DOCUMENT RESUME

ED 067 215 RC 006 483

AUTHOR	Link, Albert D.
TITLE	A Comparison of Retrieval Efficacy From the Educational Resources Information Center (ERIC) Automated Information Retrieval System Using Computer-Assisted-Instruction Training and Search Negotiations as Requester-to-System Interface Methods.
PUB DATE	Aug 72
NOTE	135p.; Doctor's dissertation submitted to New Mexico State University, Las Cruces, New Mexico
EDRS PRICE	MF-$0.65 HC-$6.58
DESCRIPTORS	*Computer Assisted Instruction; Doctoral Theses; Educational Technology; *Information Retrieval; Programed Instruction; *Search Strategies; *Telecommunication; *Training Techniques
IDENTIFIERS	*Educational Resources Information Center; ERIC

ABSTRACT

The purposes of this study were (1) to develop a requester-to-system Computer Assisted Instruction (CAI) interface with a batch processing information retrieval system, (2) to compare users' satisfaction with retrieval services between the experimental CAI training approach and the traditional person-to-person negotiation process, and (3) to test the ability of the experimental CAI interface to impart a knowledge gain of literature-searching skills related to the Educational Resources Information Center (ERIC) information base. Sixty subjects were randomly assigned to 2 treatment groups: 30 obtained computer search services by the traditional method and 30 were placed directly in charge of formulating their own search strategies via the CAI training package. Data from 58 subjects were suitable for analysis. The satisfaction scores and search precision ratios analyzed indicated that the treatment groups did not differ. However, novice users of the CAI package acquired significant gains in knowledge relative to the use of the ERIC data base. It was concluded that the use of a specially created CAI telecommunications interface system is a viable alternative to search negotiations which require the services of a professional staff. (Author/CH)

FILMED FROM BEST AVAILABLE COPY

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY

ED 067215

A COMPARISON OF RETRIEVAL EFFICACY FROM THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AUTOMATED INFORMATION RETRIEVAL SYSTEM USING COMPUTER-ASSISTED-INSTRUCTION TRAINING AND SEARCH NEGOTIATIONS AS REQUESTER-TO-SYSTEM INTERFACE METHODS

BY

ALBERT D. LINK, B.S., M.Ed.

A Dissertation submitted to the Graduate School
in partial fulfillment of the requirements
for the Degree
Doctor of Philosophy

Major Subject: Educational Administration (Research)

Minor Subject: Experimental Statistics

"PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL HAS BEEN GRANTED BY A. D. Link
TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE U.S. OFFICE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER."

New Mexico State University
Las Cruces, New Mexico

August, 1972

Copyright © 1972 by A. D. Link

CE 006483

"A Comparison of Retrieval Efficacy from the Educational Resources Information Center (ERIC) Automated Information Retrieval System Using Computer-Assisted-Instruction Training and Search Negotiation as Requester-to-System Interface Methods," a dissertation prepared by Albert D. Link in partial fulfillment of the requirements for the degree, Doctor of Philosophy, has been approved and accepted by the following:

_____

Dean of the Graduate School

_____

Chairman of the Examining Committee

_____

Date

Committee in charge:

Dr. Alfred P. Wilson, Chairman

Dr. Everett D. Edington

Dr. Timothy J. Pettibone

Dr. Darrell S. Willey

Dr. N. S. Urquhart

Dr. Gerald S. Burke

Dr. Don B. Croft

This Dissertation is Dedicated

To My Grandmother

Dr. Virginia Lee Link

. . . the most outstanding

educator I will ever know.

ACKNOWLEDGMENTS

The writer wishes to express his appreciation to Dr. Alfred P. Wilson, advisor and dissertation director, for his advice and encouragement. The writer also wishes to thank the other members of his committee for their assistance and interest in the dissertation: Dr. Everett D. Edington, Dr. Timothy J. Pettibone, Dr. Gerald Burke, Dr. Darrell S. Willey, and Dr. Don Croft. Special appreciation is extended to Dr. Scott Urquhart and Dr. Thomas H. Puckett for their many valuable technical contributions to the study.

The writer is indebted to his colleagues at ERIC/CRESS for their valuable assistance in the development of the CAI interface package (EQUIP), particularly Mr. Mike Borovicka and Mrs. Virginia Chenowith. Special recognition is extended to Mrs. Amelita Hill and Miss Manuela Quezada for their assistance in typing the first draft, and Mrs. Betty DeBellis for final typing.

VITA

Albert D. Link was born in Albuquerque, New Mexico, January 5, 1939, and resided in Bisbee, Arizona; Las Cruces, New Mexico; and Anthony, Texas, prior to establishing permanent residency in El Paso, Texas, in 1944. He is married to the former Dianne Carr of Denver City, Texas, and has two children, twin daughters Carol and Karen.

Mr. Link attended public schools in El Paso, graduating from Ysleta High School in May 1956. In August 1960, Albert Link received the Bachelor of Science degree from Sul Ross State College, and in August 1966 the Master of Education degree from The University of Texas at El Paso.

Prior to entering the New Mexico State University doctoral program, Mr. Link's public school work experience included four years as a teacher of chemistry, biology, physics, mathematics, and physical science. Additionally, he has held the positions of Evaluator of Programs (El Paso Public Schools, El Paso, Texas), Director of Research and Development, and Director of Special Programs and Services (Region XIX Education Service Center in El Paso), and served as a consultant to New Mexico State University. He has had five years' experience in the systems analysis and programming of Burroughs and IBM computers in addition to varied military training.

His professional affiliations include the National Education Association, Texas State Teachers' Association, Phi Delta Kappa, National Society for Programmed Instruction, and the American Personnel and Guidance Association.

ABSTRACT

A COMPARISON OF RETRIEVAL EFFICACY
FROM THE EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC) AUTOMATED INFORMATION RETRIEVAL
SYSTEM USING COMPUTER-ASSISTED-INSTRUCTION TRAINING
AND SEARCH NEGOTIATIONS AS REQUESTER-TO-SYSTEM INTERFACE METHODS

BY

Albert D. Link, B.S., M.Ed.
Doctor of Philosophy in Educational Administration (Research)
New Mexico State University
Las Cruces, New Mexico, 1972
Dr. A. P. Wilson, Chairman

Purposes and Hypotheses of the Study

The purposes of the study were to: (a) develop a viable requester-to-system Computer Assisted Instruction (CAI) interface with a batch processing automated information retrieval system; (b) to compare users' satisfaction with retrieval services based upon an experimental CAI and traditional search negotiation interface methods; and (c) to test the ability of the experimental CAI interface to impart a knowledge gain of literature-searching skills related to the ERIC information base. Hypotheses tested were:

(1) There are no significant differences between the measured user satisfaction with computer search results of CAI-interfaced

subjects and those undergoing search negotiations with an information specialist.

(2) There are no significant differences between the subjects' pre- and post-measures of knowledge of the information retrieval system administered during the experimental CAI treatment.

(3) There is no significant relationship between measured knowledge of the ERIC automated information retrieval system acquired as a result of CAI training and user satisfaction as measured by the user satisfaction instrument.

(4) There are no significant differences in user's satisfaction between practitioners and researchers as determined by the user satisfaction instrument.

(5) There are no significant interactions between types of users (practioner or researcher) and treatments (CAI or search negotiations) as measured by the user satisfaction instrument.

Procedures

Based upon the careful study of successful past search negotiation strategies, a totally conversational CAI training and literature search submission program named ERIC/QUERY Interface Program (EQUIP) was created taking advantage of existing APL (A Programming Language) telecommunications capabilities at New Mexico State University. Sixty subjects were assigned to two treatment groups: 30 obtained computer search services by means of search negotiation practices, and 30 were placed directly in charge of formulating their own search strategies via EQUIP. Analysis of data collected for 58 subjects was used to test the study hypotheses.

Statistical procedures included: (a) linear regression, (b) Student's t test, (c) analysis of variance, and (d) Pearson product-moment correlation.

Findings

The satisfaction scores and search precision ratios analyzed indicated that the treatment groups did not differ. It was shown that novice users of EQUIP (interfaced with QUERY, and ERIC batch processing information retrieval program) acquired significant pre- and posttest gains. Additionally, it was determined that: (a) the retrieval search precision of the two treatment groups did not differ, and (b) low search precision influenced total satisfaction scores of users of EQUIP.

Conclusions

The use of a specially created CAI telecommunications interface between uninitiated users and a complex automated information retrieval system is a viable alternative to search negotiations which require the services of a professional staff.

Table of Contents

Page

List of Tables xiii

List of Figures xiv

Chapter

1. Introduction 1

The Problem 2

The Objectives 2

The Hypotheses 3

The Basic Assumptions of the Study 4

Limitations 4

Definition of Terms 5

Organization of the Study 7

2. Review of Literature 9

Theoretical Bases 9

General systems theory 9

Information theory 9

Decision theory 10

The Use of Information Retrieval Systems 10

Batch Processing Automated Information Retrieval Systems 12

Interactive (on-line) Information Retrieval Systems . . 13

The ERIC/QUERY Information Retrieval System 15

The Educational Resources Information Center (ERIC) retrieval system 15

Related Studies 17

Chapter Page

User-to-systems interface within information retrieval programs 17

The CAI environment as a training method 18

Other studies 18

3. The ERIC/QUERY Interface Program (EQUIP) 19

The Program Language of Equip 19

The Objectives of EQUIP 20

The Development of EQUIP 21

A Generalized Description of EQUIP 22

Block one: system familiarization 23

Block two: QUERY familiarization 23

Block three: search strategies 23

Block four: practice and submission 23

The Instructional Sequence of EQUIP 28

Discussion 28

4. Procedures 33

Design of the Study 33

Setting of the Study 35

Summary of the Experimental Design 36

Sampling Design 36

The study population 36

Sampling technique 38

Limits of inferences to the study population 39

Measures Employed 39

Pre- and posttesting of EQUIP 40

Chapter Page

Implications 71

The use of EQUIP 71

Utilization of user characteristics 72

Recommendations 72

References 74

Appendix A. EQUIP 78

Appendix B. Sample Session 166

Appendix C. Instruments and Forms 203

Appendix D. Technical Notes 207

List of Tables

Table		Page
1.	A Contrast Between DIALOG and QUERY Information Retrieval Systems	16
2.	Subblocks and Objectives of Block 1 of EQUIP	24
3.	Subblocks and Objectives of Block 2 of EQUIP	25
4.	Subblocks and Objectives of Block 3 of EQUIP	26
5.	Subblocks and Objectives of Block 4 of EQUIP	27
6.	Frequencies of Study Sample Subjects in Relation to Retrieval Expectations	45
7.	Study Sample by Treatments	47
8.	Numbers and Percentage of Various Categories of Subjects Within Treatment Groups	48
9.	Summary of Statistical Computations Based on User Satisfaction Scores	51
10.	Summary Table for Experimental and Control Treatment Groups Based on Total Satisfaction Scores Obtained on the User Satisfaction Instrument	53
11.	AOV for Testing Hypotheses One, Four, and Five	55

List of Figures

Figure		Page
1.	Major Instructional Sequence of EQUIP	29
2.	An Ideal Research Design	33
3.	The Experimental Research Design	34
4.	Basic Design of Experiment	37
5.	Summary Calculations of $\bar{t}'$ Statistic for Item Three of the Satisfaction Questionnaire	50
6.	Summary of Calculations Comparing the Pre- and Posttest Interface Knowledge Gains of the Experimental Treatment Group	56
7.	Summary of Linear Regression Computational Results Relating to the Predictive Relationship of Posttest Scores and Pre- and Posttest Gains to Experimental Subject's Satisfaction Scores	58
8.	Scattergram of Experimental Group Total Satisfaction Scores Versus Pre- and Posttest Gains	60
9.	A Contrast and Correlation of Experimental Total Satisfaction Scores and Search Precision for Cluster Two	61
10.	A Contrast of Linear Regression Summary Calculations for the Cluster 1 and Cluster 2 Subjects Shown on Figure 8 . . .	62
11.	Summary of Statistical Computation for Determining the Significance of Experimental Group Precision Means . . .	64

Chapter 1: Introduction

Information is being produced by all disciplines at an increasing rate each year. In response to this phenomenon, the United States Office of Education (USOE) has committed millions of dollars of the nation's resources to the creation of the first national education information system, known as "ERIC" (Educational Resources Information Center), to assist the decision-maker, researcher, and practitioner assess a growing bank of documents for the education profession.

To aid the serious user in his document search and retrieval tasks, in 1969 a computerized searching system known as QUERY was created and implemented in over a dozen installations across the nation. QUERY is a batch processing system which requires a unique "search language" as machine-readable input and offers almost unlimited document searching capabilities. Because of this distinctive characteristic, if one is to properly use the QUERY automated information system he must be able to select the correct search criterion prior to submitting his search request to the computer.

Currently, all QUERY installations accommodate the information seeker (user) by negotiating a search request. This requires a conference between the user and an information specialist and is concluded when an acceptable search strategy (search language) based upon the user's expressed needs is recognized by the information specialist. This method of interfacing the system (through the information specialist) is less than satisfactory for a number of

reasons including: (a) the difficulty the user has in expressing his specific search need to another person; (b) the size of the files often "overpowers" the unitiated user of the ERIC system; and (c) the search process (via QUERY) is not completely explained to the user, thus his knowledge of the retrieval capabilities is severely limited.

The Problem

Because of existing difficulties experienced by uninitiated users of the ERIC automated information retrieval system, which currently depends upon user interface with the system via information specialists, new approaches allowing the potential user of ERIC files to meet his information needs must be explored. For the system to have maximal impact in the education sector, a more efficient interface methodology must be created that will give the potential user opportunities to receive relevant experience and training allowing direct interface with the information base.

The Objectives

The objectives of this study were:

(1) To develop, implement, and test a fully documented computer-assisted-instruction (CAI) package to train potential users of the computerized ERIC files to interface with the information retrieval system without the assistance of an information specialist. Effectiveness of the CAI program as a training package will have been assessed through the use of pre- and posttesting.

(2) To design an instrument to measure the system user's satisfaction with the results of computer searches from either the information specialist or from personal interface with

the QUERY information retrieval system.

(3) To determine any differences between experimental and control groups of ERIC automated information retrieval system users (including predesignated subgroups of each)--with the experimental group undergoing CAI training on the use of the system and the control group receiving no such training--through the use of a user satisfaction instrument and appropriate statistical analysis.

The Hypotheses

The specific hypotheses tested were:

(1) There are no significant differences between the measured user satisfaction with computer search results of CAI-interfaced subjects and those undergoing search negotiations with an information specialist.

(2) There are no significant differences between the subjects' pre- and post-measures of knowledge of the information retrieval system administered during the experimental CAI treatment.

(3) There is no significant relationship between measured knowledge of the ERIC automated information retrieval system acquired as a result of CAI training and user satisfaction as measured by the user satisfaction instrument.

(4) There are no significant differences in user satisfaction between practitioners and researchers as determined by the user satisfaction instrument.

(5) There are no significant interactions between types of users (practitioner or researcher) and treatments (CAI or search negotiations) as measured by the user satisfaction instrument.

The Basic Assumptions of the Study

The study is based on the following assumptions:

(1) Once an educator understands the utility of an information base especially designed and maintained for his discipline, he will attempt to capitalize upon the respective benefits of such a system.

(2) Potential user traits and skills will vary extensively; reading ability, typing ability, and educational level (i.e. educators vs. trainees) will differ.

(3) The ERIC information file size will not remain static; file size currently increases over one thousand documents per month (file size was held constant for this study).

(4) For the duration of the study, computer files, search programs, and listing formats will be unchanged.

(5) Measurement instruments are valid.

(6) Subjects have had no prior experience with ERIC automated retrieval systems.

Additionally, it should be noted that the QUERY information retrieval system is assumed to exist in a stable form. The specific version of QUERY used was not altered in any manner for the duration of the study. In this manner, the interface system used was the primary dependent variable. Accordingly, QUERY was not an object of interest for this study.

Limitations

The systems environment of the New Mexico State University (NMSU) Computer Center is not unlike that of other installations using the QUERY information retrieval software. There are, however, some

004

unique systems capabilities which are not in existence at many ERIC installations. For example, not all ERIC installations possess APL (A Programming Language) capabilities, and of those that do, not many have the ability to initiate a search from the APL subsystem to the main operating job stream. It is this capability which allows for a true interface of the user to the batch processed QUERY system. To be more specific, this study is highly dependent on computer capabilities--some of which could be installed at other computer sites in the near future.

Definition of Terms

APL. "A Programming Language" is an IBM-supported telecommunications language. The version used for this study is APL/360-05, Program Number 5736-XM6, modified for data-to-job-stream capabilities by Dr. Thomas H. Puckett, NMSU Computer Center.

Automatic information retrieval. The term "automatic information retrieval" refers to any viable computer-based information retrieval system having document isolation and retrieval capabilities using inverted or sequential files.

Batch processing. In contrast to on-line processing, batch processing is a technique by which the items to be processed in a data processing machine must be collected into groups prior to processing (Williams, 1965, p. 402).

Boolean. Algebraic operations on, and functions constructed from, two-valued variables, interpreted as representing truth-falsity or set membership status is a characteristic of Boolean logic (Williams, 1965, p. 403).

CAI. "Computer Assisted Instruction" is a technique which capitalizes upon the capabilities of the computer to store and selectively provide information to a learner through the use of a telecommunicator device.

CIJE. Current Index to Journals in Education is a file comprised of citations from more than 500 major educational publications: journals, quarterlies, annuals, and yearbooks.

Descriptor. Any word or key word used to classify a document for subsequent retrieval is designated a descriptor.

File. A file refers to a structure of information composed of one or more records, such that the records are descriptive of individual document citations or classes of citations.

Hit. A hit is a retrieved document citation (listing) which is considered relevant to an information need stated by the user.

Interface. A method or technique of linking working systems without deteriorating the systems involved is known as an interface.

On-line retrieval system. This term refers to an automated information retrieval system which allows the user access to specific sections of an information file via random access techniques and provides immediate feedback relating to his search; contrasted with batch processed information retrieval.

Practitioner. For the purposes of this study, a seeker of information who desires a small number of highly relevant hits as a result of his search effort is classified as a practitioner.

Researcher. A researcher is a seeker of information who desires a large number of relevant and slightly relevant hits, or

who expects a search to retrieve a considerable number of document citations with emphasis on "in-depth" criteria.

RIE. Research in Education is a file comprised of completed research and research-related reports in education as input by the various ERIC clearinghouses.

Search language. A specific encoded Boolean search format used to access citations with an automated information retrieval system, such as QUERY, is called a search language.

Search negotiation. Search negotiation is a process whereby a user communicates his information need to an ERIC information specialist who, in turn, encodes the necessary computer input.

Search precision. Search precision is a percentage ratio of the number of hits a search produces to the total number of document citations retrieved.

Search request. A single search encoded in an acceptable search language for QUERY which has been submitted to the computer job stream is known as a search request.

Terminal. An IBM 2741 telecommunication device equipped with an APL keyboard and type head (font ball) is known as a terminal.

User. A user is one who seeks information through the use of an automated information retrieval system (QUERY).

Organization of the Study

Chapter 2 presents a review of available relevant literature and related studies which focus on the problem. Chapter 3 provides the major characteristic of EQUIP (ERIC/QUERY Interface Program), the experimental treatment considered by this study. A description of

the procedures used is found in Chapter 4; the analysis of data is given in Chapter 5; and findings and conclusions are reviewed in Chapter 6. Of possible interest to future implementators of EQUIP, the entire encoded package is provided in the Appendix as well as technical notes and a sample of an EQUIP sample session.

Chapter 2: Review of Literature

Of the four sections to this review, the first considers the theoretical bases of the problem to be studied, the second deals briefly with information retrieval, the third concentrates on the ERIC/QUERY information retrieval effort, and the last focuses on related studies.

Theoretical Bases

General systems theory. General systems theory as described by Optner (1960, p. 9) and Stufflebeam (1971, p. 124) is concerned with two major characteristics--input to the system and output from the system. When such systems involve computer usage they ". . . are structured, or designed, to operate in nonvariant, highly predictable ways [Optner, 1960, p. 4]." It follows that if a highly structured system is not understood by its users, the efficacy of such a system will be less than design expectations.

This study attempted to show that, within the realm of information retrieval systems, the less knowledge the user of the system possesses about the system, the greater the chances the system will not operate within its designed purpose.

Information theory. Shannon (1959, p. 18) indicates that a general communication system is comprised of an information source which originates a message (input), a transmitter, a channel and a receiver which processes the message, and a destination (output). Meadow (1967, p. 6) renames Shannon's intermediate processing steps (transmitter, channel, and receiver) by referring to them as

"transducers" having the capability of selection as well as processing. This, according to Meadow, allows for message processing through time with less regard for knowing the identity of the originator or the destination. Thus, it seems, the merging of the computer system and the information system allows for input (documents) and output (retrieved documentation) within the context of both general systems and information systems.

Decision theory. It is interesting to note that the various decision-making theories of Kepner and Tregoe (1953, p. 73), Griffiths (1959, p. 90), Halpin (1966, p. 35), and Meyer (1968 vol. 1) each rely on the acquisition of information as a prelude to completing the decision-making process. Additionally, each process is described as a system or may be thought of in terms of general systems theory. Stufflebeam (1971, p. 38) points out that the availability of information prior to selecting alternatives in decision-making is assumed. Thus, the importance of information within the context of educational decision-making may not be overlooked.

010

The Use of Information Retrieval Systems

Lancaster (1968, p.1), who has been associated with the Medical Literature Analysis and Retrieval System (MEDLARS) states: "An information retrieval system does not inform (i.e., change the knowledge of) the user on the subject of his inquiry. It merely informs him of the existence (or nonexistence) and whereabouts of documents relating to his request." This suggests that such a system may consist of the simplest card catalogue or filing procedure

to the highly sophisticated computer-based automated systems which exist today. Salton (1968) describes such an automated system, called SMART, which has been installed at Harvard and Cornell Universities. Both Williams (1965, chap. 10) and Lancaster (1968, chaps. 1-3) describe traditional and modern information retrieval schemes. Noteworthy in these discussions is the large number of variations available to the designers of these systems.

The use of an automated retrieval system is anything but a simple matter. Lancaster (1968 , pp. 181-2) suggests that retrospective literature searches may be divided into those conducted without an intermediary by the person having the information need, and those delegated by this person to a second individual, usually a librarian or information specialist. Lancaster (1968, pp. 182-5) clearly shows that there are different skills and expectations required of the requester under these two schemes. Additionally, he points out that twenty per cent of the National Library of Medicine's Medical Literature Analysis and Retrieval System's searches involving defective interaction (between the user and the searcher) were judged to be of the type in which the requester, using an intermediary, was unable to precisely define his need except through some browsing in the literature (1968, p. 184). To summarize, this means that to be successful the requester must spend a considerable amount of effort defining and negotiating his search request with the person who will actually encode the search in machine-readable form or learn how to use the system himself. The economics of the former appear to be less attractive than those of

011

the latter if it is assumed that the same effectiveness (in terms of precision of retrieved information) can be achieved.

Batch Processing Automated Information Retrieval Systems

In the fourteen years since the first published application to key word indexing of precoordinated information files by the computer (Janda, 1968, p. 4), hundreds of batch processing information retrieval systems have been created. Janda (1968) describes a number of systems receiving attention within the newly emerging discipline of information retrieval. These include: KWIC (Key-Word-In-Context), which uses Ceagaard's BIDAP (BIbiographic DAta Processor) computer program for input processing; TRIAL (Technique to Retrieve Information from Abstracts of Literature); and SEARCH (Janda, 1968, p. 28), which uses magnetic tapes and a simple search scheme.

Adams, writing in a collection of papers (Rubinoff, 1965, pp. 77-85), describes MEDLARS, which has some similarities to the ERIC/QUERY system briefly described in the third section of this chapter.

A review of the literature shows that many batch processing information retrieval systems were either designed for highly specific applications and/or developed around available computer hardware and software.

The apparent advantages of the batch processing systems include:

(1) Systems designers may capitalize on their existing computer capabilities without added costs for additional software.

(2) Maximal specialization of the system may be facilitated, allowing for "tailor-made" systems which meet the designer's needs at the least cost.

(3) "In-house" created computer programs are maintained internally; thus changes to the system may be made with the fewest constraints.

Because of the advantages listed, it is unlikely that batch processing information retrieval systems will be discontinued without some significant changes in the costs and "state of the art" of the computer industry.

U13

It is noteworthy that, of all of the batch processing information retrieval systems investigated by the author, none were accompanied with a system-to-user interface. All required the intervention of an intermediary, as cited by Lancaster (1968, pp. 182-5).

Interactive (on-line) Information Retrieval Systems

Caruso (1969, pp. 7-18) describes eight on-line interactive retrieval packages, including Lockheed's DIALOG system which will be discussed briefly. In their independent evaluation of the DIALOG system, Timbie and Coombs (1969, pp. 21-31) present their subjects' comments in the form of good and bad features. The favorable features of DIALOG were:

> The system is thorough, comprehensive.
> The system is effective.
> The system is generally valuable.
> The system is exhilarating, interest-creating.
> Having a hardcopy record to refer to is handy.

Unfavorable features were:

> Too many combinations of keys are needed to input one command,
>
> Having to build combined sets one step at a time, rather than using parentheses, and doing it with one complex statement is inconvenient, and
>
> There is a great deal of "paging" required on the CRT, because you can only look at nine terms at a time.

Additionally, two other characteristics of the DIALOG system are suggested: speed, and "horizon widening"--the capability of the system to expand or limit a search (kin to browsing).

The Timbie and Coombs evaluation is based on nine case histories and is not a controlled evaluation of the system. Their evaluation does suggest that interactive on-line systems such as DIALOG will meet ERIC users' needs. The estimated cost provided by Lockheed for a "typical" regular DIALOG user is $35,000 per year for twenty hours of searching per week (Lockheed, 1971b), a cost which is relatively high for most educational institutions.

A. R. Barnum (Rubinoff, 1965, pp. 45-61) describes the Air Force's Reliability Centered Data Management System (RCDMS) which has on-line retrieval and analysis capabilities. Inherent in the on-line systems found in the literature is the capability of allowing the user to browse through the information file while using immediate feedback--in short, to interact with the file. Some on-line systems such as GRINS and EASY ENGLISH (Caruso, 1969, p. 10) use English-language statements as a search criteria, while others such as the Lockheed Company's DIALOG (1971a, p. 4) use a Boolean search strategy similiar to that of QUERY. Since a user

014

must be individually trained to use the system, most interactive (on-line) systems provide a users' training manual or, as was the case with Caruso's study (1969), provide CAI tutorial training.

There exist only two nationally used automated information systems designed to access the ERIC information base, QUERY and Lockheed's DIALOG. Table 1 presents a contrast of these two systems based on Lockheed's (1971a) user manual and Computer Resources Corporation's (1970) installation manual, and ERIC/CRESS experiences.

The ERIC/QUERY Information Retrieval System

The Educational Resources Information Center (ERIC) retrieval system. The ERIC system was begun in 1965 with the production and dissemination of the Catalog of Selected Documents on the Disadvantaged, and since has grown into an international educational information retrieval system of considerable scope and complexity (Burchinal, 1970, pp. 58-63). Currently, ERIC files may be searched manually or via the computer using an on-line system called DIALOG or batch processing programs such as the USOE-sponsored QUERY system.

ERIC depends upon the collective efforts of 17 clearinghouses located in various areas of the country for the acquisition of relevant documents into the ERIC information base. One such clearinghouse, ERIC/CRESS (Clearinghouse on Rural Education and Small Schools), is located on the New Mexico State University campus. In 1969, the USOE-sponsored automated information retrieval program, QUERY, was installed at ERIC/CRESS. For the first time, educators in the Southwest had at their disposal a powerful information retrieval tool to assist them in their professional activities.

Table 1

A Contrast Between DIALOG and QUERY Information Retrieval Systems

Characteristic	DIALOG	QUERY
Search Input	Structured	Structured
Feedback	Yes	No
Thesaurally Related Terms	Yes	Yes (Manuals)
Count of Citations Retrieved	Yes	Yes
Logic	Boolean	Boolean
Output Format	Citation	Citation
User-to-system Interface	Yes (Instruction Manual)	No
Cost (overall, per year)	High ($21,000)	Moderate (est. $8,000)
Turnaround Time	1-5 days (mailed)	1-3 days
Type of Processing	On-line (interactive)	Batched
Communication Mode	Terminal or CRT (Cathode Ray Tube)	Card Reader (Punched Cards)
Browsing Allowed	Yes	No
Computer Facilities	Centralized	Decentralized

U16

QUERY is an advanced, generalized, sequential, file-searching computer software package, using RIE and CIJE magnetic tapes or disk files, which is capable of isolating and listing any informational subset, depending on the search strategy used (Computer Resources Corporation, 1970). In late 1970, David Altus, while a graduate assistant at ERIC/CRESS, modified the QUERY search subprogram to expand the power of the search language. Mr. Altus' change to QUERY allows the use of a versatile TEXT modifier--a function which allows any part of a word, phrase, or sentence to be scanned during the record-by-record search operation of QUERY.

Brandhorst and Marra (1970) provide a comprehensive description of RIE and CIJE magnetic tape files produced by the Leasco Company. A brief file description is found in appendix D.

The evaluation of the ERIC system as an information retrieval system as described above is not considered in this study. Rather, the use of the system in relation to the current practice of providing an information specialist as an intermediary versus the alternative of training the requester to formulate his own search strategies is pivotal. The computer files (RIE and CIJE) and programs (QUERY) are independent variables.

Related Studies

User-to-systems interface within information retrieval programs. Caruso (1969) found no significant (.05 level) differences between an experimental group of Masters-level library science students trained to operate an on-line information retrieval system using a computer-assisted tutorial program and a control group trained by class

lectures. This seems to suggest that subjects may be trained using a computer terminal just as effectively as more costly lecture classes or individual person-to-person tutoring. Also, Caruso's study was done using an on-line retrieval system--not a batch processing system such as QUERY--a difference which is not critical in terms of training, but which does leave the interface method question completely unanswered.

The CAI environment as a training method. Sutter (1967) found no achievement differences between groups of subjects working alone with a CAI device (terminal) and those working with the device assisted by an individual tutor. Assuming an effective CAI program may be created to train requesters to use the ERIC information retrieval system (computerized), Sutter's findings seem to suggest that there will be no need to provide individual tutors to assist in the training process. Combining the results of both Caruso and Sutter, within the scope of this study, suggests that the use of existing computer terminals at New Mexico State University to explore the problem is legitimate.

Other studies. Although the literature abounds with studies, projects, and programs which use CAI as a training media, no CAI programs or studies are known to exist which specifically focus upon the problem of requester-to-system interface.

ERIC

Chapter 3: The ERIC/QUERY Interface Program (EQUIP)

Available literature clearly shows the existence of numerous batch processing information retrieval systems as well as an increasing number of on-line interactive systems. The investment of resources for batch processing systems will most likely insure their existence and maintenance for years to come. The on-line systems are gaining in popularity despite their cost, due in part to the ability to provide systems-to-users interface, a feature totally lacking with the more numerous, less costly batch processing systems. This phenomenon led to the present study. If it is possible to create a satisfactory interface package to be coupled to a relatively complex QUERY batch processing computer package, greater efficacy of the system's use should result.

Since no search interface package was known to exist, the writer created EQUIP (ERIC/QUERY Interface Program), a CAI training and search submission package described in this chapter.

The Program Language of EQUIP

IBM's popular telecommunication language, APL, available at New Mexico State University, was selected as the language to program EQUIP because of its immediate availability, versatility, and power as a CAI medium. The version of APL used is not standard, as supported by IBM, but instead has a unique feature provided through the efforts of Thomas H. Puckett of the NMSU Computer Center. Dr. Puckett's modification of APL allows for the submission of the searches encoded by the user directly to the job stream of the

university's IBM 360/65 multiprocessing computer job stream--a feature not currently available in the IBM version of the language.

In addition, APL has been found to be efficient as related to computer time used for interface purposes. This characteristic coupled with some unique matrix operations and generic time-monitoring functions led to the decision to use APL as the CAI program language for EQUIP.

The Objectives of EQUIP

The second and most important consideration relating to the creation of EQUIP was to determine a sound training foundation and minimum instructional objectives. This process relied heavily on the writer's two years of experience while maintaining the QUERY system and processing QUERY searches for hundreds of users. Thus the central focus used in deriving minimal performance objectives for EQUIP are centered on the criteria: What is the minimum a novice user must learn for the creation of a satisfactory computer search? Based on experiences with the system, the following guidelines were used:

(1) The user must understand the proper use and meaning of three standard reference sources--the ERIC Thesaurus of Descriptors, the ERIC Rotated Thesaurus of Descriptors, and the ERIC Posting of Descriptor Statistics (a source of the number of times a particular descriptor is used by RIE).

(2) The user must be able to create a statement of his information need and subsequently reduce that statement to basic elemental descriptors.

(3) The user must be able to properly use descriptors to isolate the citations of interest from the ERIC information files.

(4) The user must understand and be able to correctly use the basic Boolean logic required in the QUERY search language.

(5) The user must be able to correctly encode a search in machine-readable language.

These objectives were determined to be the minimal learning objectives for EQUIP. Since it was desired that the user-to-system interface program be suitable to individual differences (via CAI), other less important objectives were evident:

(1) The user must learn how to use the telecommunication terminal.

(2) The user must be aware of what the computer (QUERY) can and cannot do.

The Development of EQUIP

Based on the previously defined criteria, EQUIP was planned and written. Each subprogram was written and tested using both knowledgeable and novice subjects until the specific objectives of the subprogram could be met to the satisfaction of the writer 100 per cent of the time. As would be imagined, some subprograms were rewritten several times before they were deemed acceptable.

After subprograms and related support programs were written, tested, and accepted, the total package was assembled and tested using knowledgeable and novice users of ERIC/QUERY facilities. Necessary modifications were installed and "polished" until EQUIP was considered ready for formal testing and comparison with existing

search negotiations procedures.

An integral part of EQUIP is an eighteen-item pre- and post-test administered to the user in the course of his CAI training and subsequent search submission. (These measures are discussed more fully in Chapter 4.) The prime purpose for the pretest was that of switching (branching) the training sequence of the user with EQUIP, based on his knowledge of the ERIC system. It was assumed from the beginning that the skills and knowledge of users would vary widely. Through the use of the pretest switching provision of EQUIP, the training a particular user might undergo was, essentially, individually prescribed based on what "he brings to the training session." It should be noted that the only major branching with EQUIP takes place as a result of the user's performance on the pretest.

A Generalized Description of EQUIP

EQUIP[1] is comprised of four major instructional blocks, each consisting of subblocks or units. It should be remembered that, based on user's knowledge of ERIC or other criteria (familiarity with the APL keyboard, etc.), it is possible that some users received no exposure to some blocks or subblocks of the training program. The four major blocks and respective objectives are discussed.

---

[1]No attempt will be made to present a comprehensive description of EQUIP here. Interested readers will find a complete copy of the APL coded programs (functions) in Appendix A and a sample of an EQUIP session in Appendix B.

Block one: system familiarization. The major purpose of this instructional block is to allow the user to learn the basic use of the APL telecommunication medium and some very basic concepts unique to computer systems. Internal subblocks and objectives are shown in Table 2.

Block two: QUERY familiarization. The second major instructional block concentrates on the standard published materials deemed necessary to assist the user in his search construction. Table 3 depicts the subblocks and objectives of the second instructional block.

Block three: search strategies. Only after the user has received appropriate training or has shown that he is knowledgeable of the concepts and materials offered in the first two blocks may he embark on block three, the most critical portion of the instructional sequence. Block three deals exclusively with the construction of the user's particular search strategy of interest. Subblocks and respective objectives are given in Table 4.

Block four: practice and submission. The fourth and final block of instruction provides the user unlimited practice in creating his actual search, encoding it, and, subsequently, submitting it as a job to the computer job stream. Additionally, a posttest is administered, scored, and recorded at the end of this block. Table 5 presents the subblocks and instructional objectives of block 4.

Table 2

Subblocks and Objectives of Block 1 of EQUIP

Subblocks	Objectives
Terminal Training	The user must demonstrate competency in using the IBM 2741 telecommunication terminal.
Keyboard Training	The user must demonstrate competency in using specific APL keyboard characters as well as knowledge of basic keyboard functions such as correcting mistakes.
Response Training	The user must show competency in responding in the correct manner to instructions, questions, or commands used within EQUIP.
Pretest	A means of ascertaining the user's knowledge base relative to the QUERY automated information retrieval system must be provided.
Systems Training	The user must demonstrate a basic knowledge of simple computer input, output, and processing concepts.
Computer Input Familiarization	The user is shown how the general search language (encoded) appears as input to QUERY.
Computer Output Familiarization	The user is invited to investigate sample listings of citations provided near the terminal.

Note.--With the exception of the last two subblocks, the user must demonstrate competencies prior to being advanced further in the session.

Table 3

Subblocks and Objectives of Block 2 of EQUIP

Subblocks	Objectives
File Familiarization	The user must show an acceptable knowledge of the use of existing ERIC files and the relationship and usage of descriptors in isolating documents from a file.
Rotated Thesaurus Familiarization	The user must indicate that he has learned the use of the ERIC rotated thesaurus and its application to modified descriptors.
Descriptor Statistics Usage	The user must demonstrate knowledge and use of the ERIC descriptor reports, a resource material which aides in predicting the number of possible search hits.
Thesaurus Usage and Familiarization	The user must demonstrate the ability to use descriptors, as given in the ERIC thesaurus, in relation to his information need statement.

Note.--Due to the wide variability of possible correct user responses, progress is monitored through the use of answers given to specific questions which are used to switch (skip over) certain instructional material. Progress is not always performance-based as was noted in the first block of instruction.

Table 4

Subblocks and Objectives of Block 3 of EQUIP

Subblocks	Objectives
Selecting Descriptors	Using previously learned concepts, the user tentatively selects the descriptors he actually intends to use for his search strategy.
Logic Training	Using preselected descriptors from the previous subblock, all possible allowed logic sequences must be presented and understood by the user. It is at this point in the instruction that the user must indicate his understanding of the Boolean logic used by QUERY.
Estimating Hits Familiarization	The user must be cognizant of the possible list his set of descriptors will produce, given all possible Boolean logic operator combinations. If the user is not satisfied, he must have the option of either changing his descriptors or selecting a different logic pattern.

Table 5

Subblocks and Objectives of Block 4 of EQUIP

Subblocks	Objectives
Encoding the Search	The user must be given the opportunity to be actively involved in the encoding of his search.
Diagnostics Rework and Submission	Mechanical encoding errors must be detected and feedback given to the user as he learns to encode his search. The user must demonstrate competency in encoding his specific search of interest. The user's search request must be rejected by EQUIP if it fails the search language criteria of QUERY. If the strategy is acceptable, the user must be allowed to submit his search directly to the computer job stream.
Posttest	A means of ascertaining the user's newly acquired knowledge of automated information retrieval via QUERY and EQUIP must be accomplished

Note.--The posttest appearing as the last subblock will be deleted from EQUIP after final testing of the package.

The Instructional Sequence of EQUIP

EQUIP utilizes a flexible instruction sequence based on initial switching as a result of user responses to the eighteen pretest items provided in block one. A simplified flowchart of the sequence decision system is presented in Figure 1. It should be noted that not all switching of the sequence is based on pretest responses. Additional switching is accomplished via binary choice questions requiring honest user responses. For example, if asked "Do you know how to use the rotated thesaurus?" the user must respond "No" to receive needed training. Work space limitations prevented total competency-based testing of the user's knowledge--thus the "question and response" technique of controlling instruction sequence was necessary at times.

Discussion

The cumulative elapsed time of instructional block and subblock of EQUIP is recorded on two separate time clocks--real time and unlocked keyboard wait time (terminal waiting for user responses). If the user elects to undergo retraining, when offered, his instructional times are accumulated with those of his first experience with the subblock concerned. In this manner, not only a record of times was recorded for each user but the blocks and subblocks actually used in the training sequence were easily identifiable.

Separate item-by-item records were kept of the user's responses to both pre- and posttest sessions. These were recorded binarily--zero for an item answered incorrectly and one if answered correctly.

Figure 1

Major Instructional Sequence of EQUIP

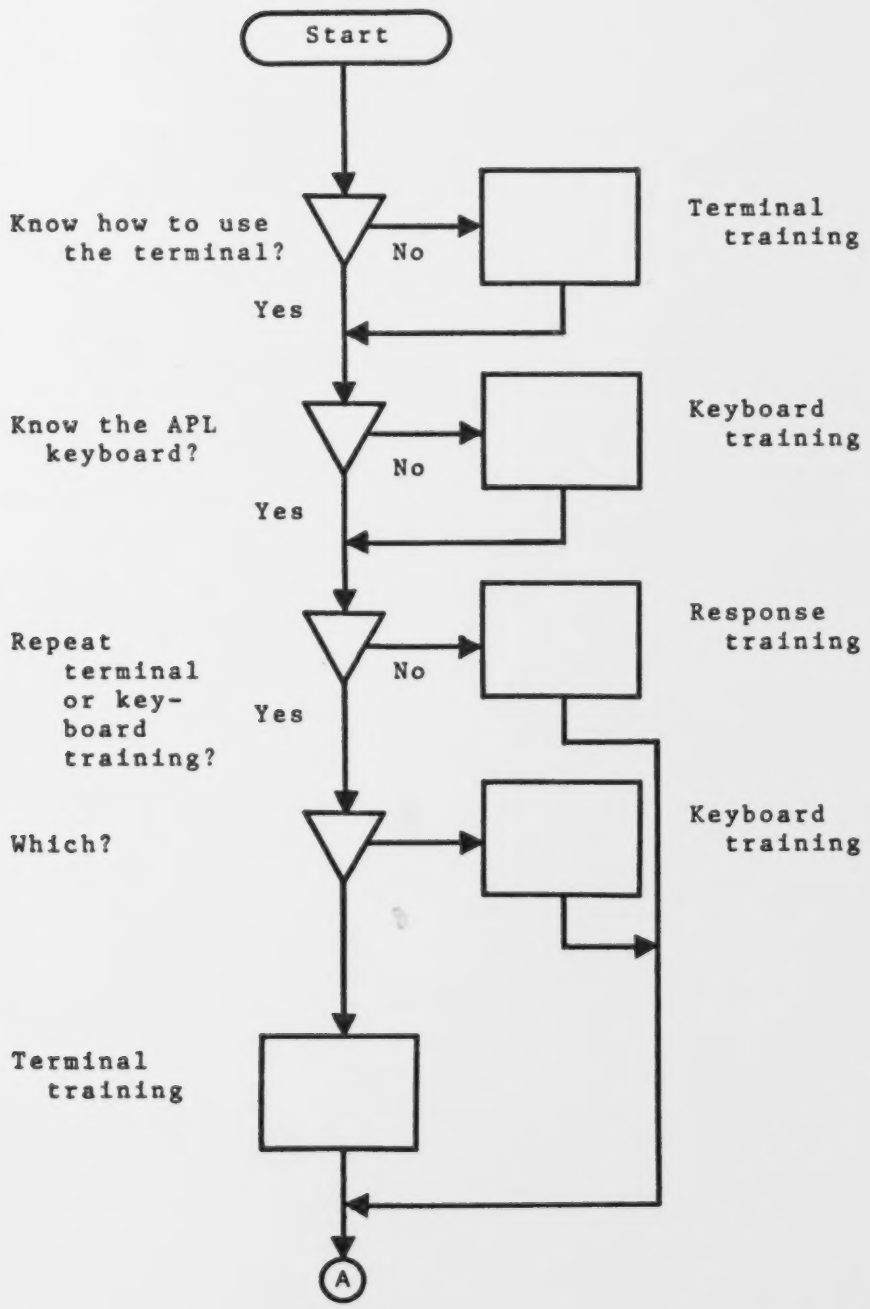


Figure 1 (Contd.)

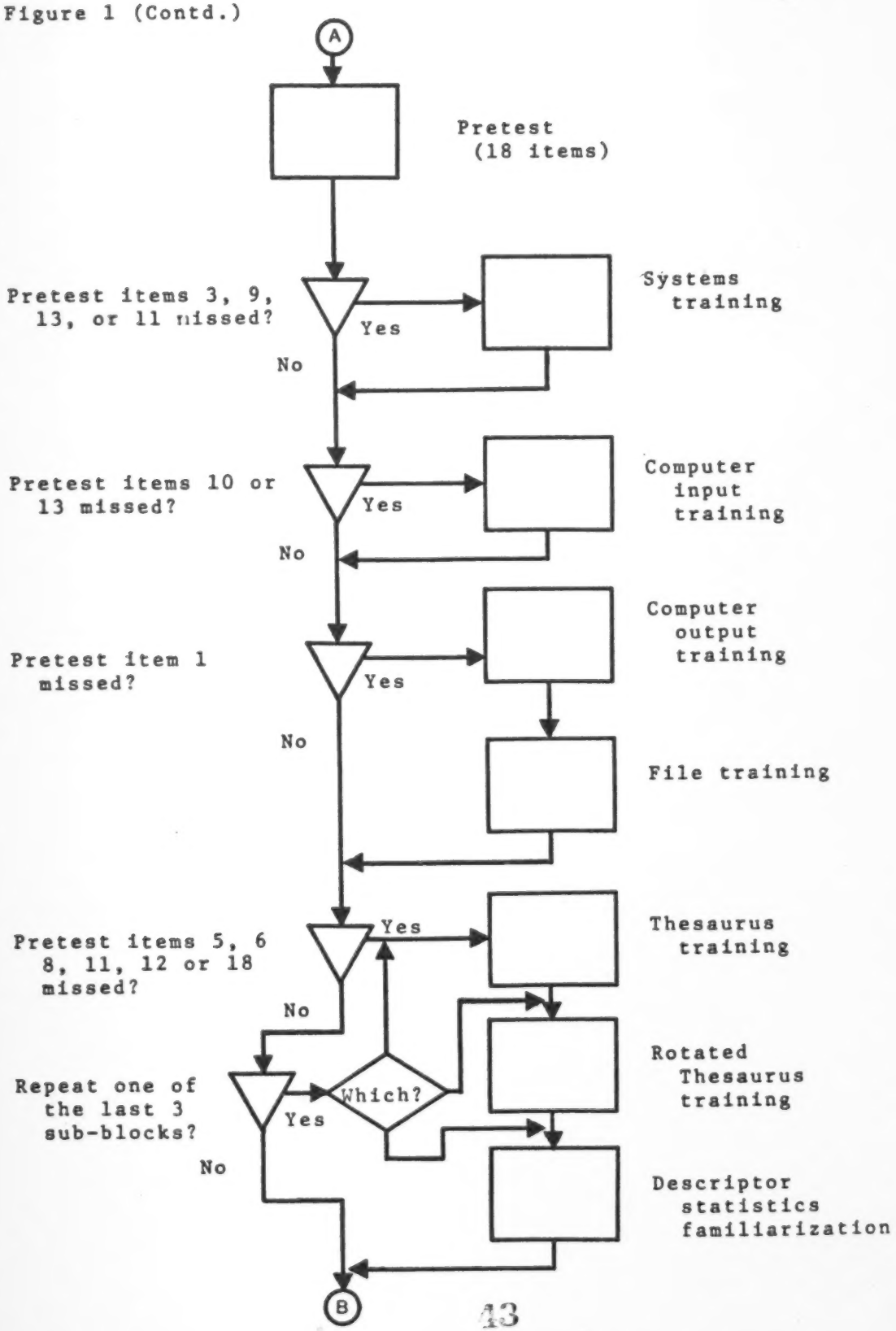
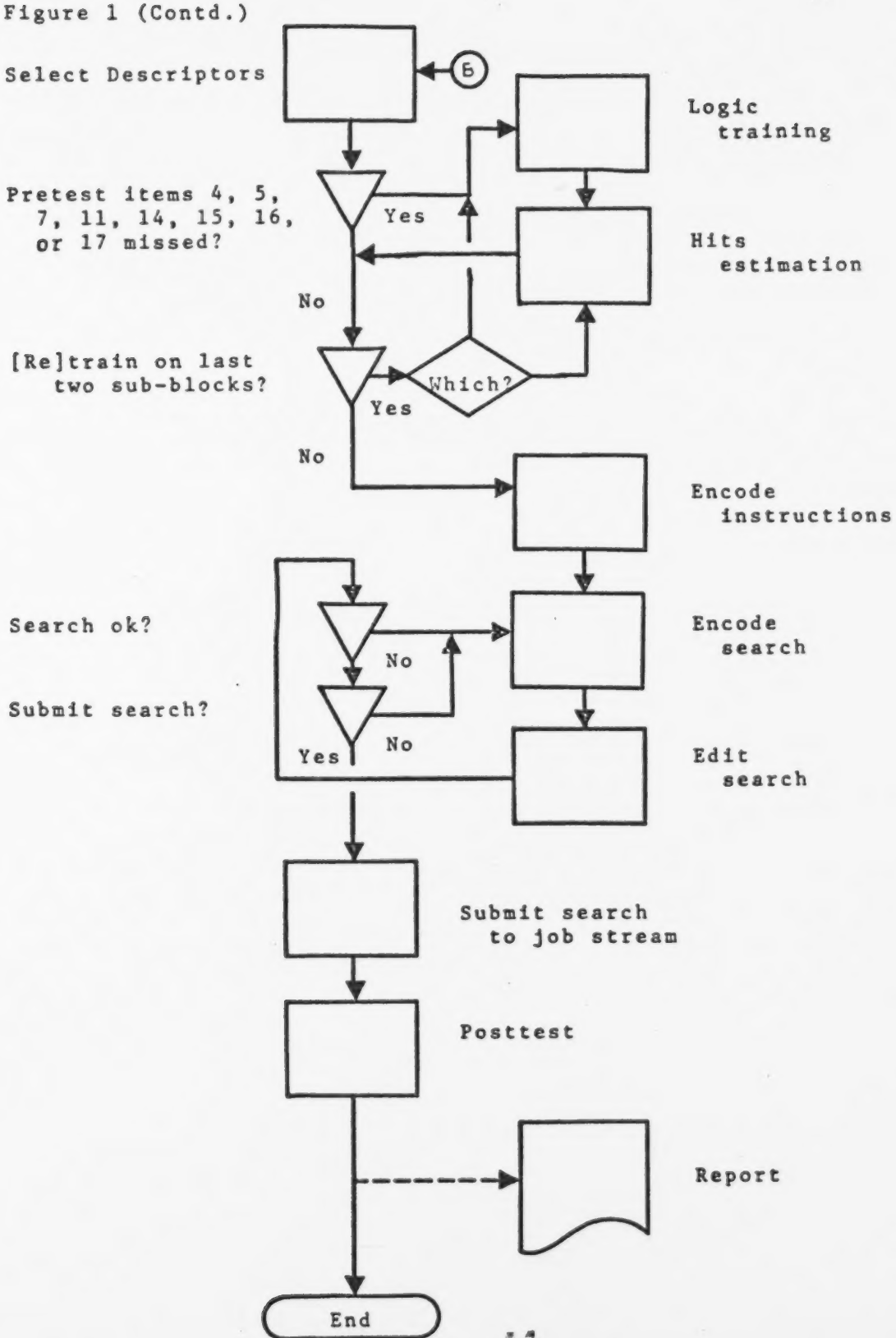


Figure 1 (Contd.)



A master real-time clock recorded the total training session elapsed time and another clock was employed to keep a cumulative record of CPU time used from sign-on to completion of the session.

A sample copy of clocked times and other monitored variables may be found in the last pages of Appendix B.

Chapter 4: Procedures

Design of the Study

The rational of this study, as developed in the first two chapters, reduces to the following experimental question: After a CAI interface package has been created, is there any empirical evidence that its use with persons desiring information from the ERIC information base is effective? To phrase this question in another way: Is EQUIP as good an interface technique as existing search negotiation sessions with an information specialist?

Borg and Gall (1971, pp. 363-401) suggest a number of experimental research designs of increasing complexity. An ideal design, they suggest, is one which offers the maximal control of the experiment by the investigator. A suggested "acceptable" design for educational research is diagramed below (Borg and Gall, 1971, p. 376).

R	O_1	X_c	O_2	(control group)
R	O_1	X_E	O_2	(experimental group)

Figure 2. An Ideal Research Design

Briefly, Figure 2 represents a "classical" experimental design where subjects in both the experimental and control groups are randomly chosen from the population, randomly assigned (R) to the treatment groups, pretested (O_1), administered an experimental (X_E) or control treatment (X_c), and posttested (O_2).

Unfortunately, the nature of this study did not lend itself

to this highly desirable design structure. Instead, a design which offers less control of the experimental situation was required, since it was impossible to predetermine a randomly chosen sample of subjects from the population of potential users of ERIC/CRESS facilities. The offering of computer searches is, by nature, service oriented. The subjects, therefore, either do or do not desire services--a decision made on an individual basis. This characteristic, coupled with the fact that the search negotiation interface technique requires no prior knowledge of the ERIC system, led to the use of the following experimental design.

R		X_c		O_2	(control treatment)
R	IP_1	X_E	IP_2	O_2	(experimental treatment)

Figure 3. The Experimental Research Design

As shown in Figure 3, subjects are randomly selected from the population (R), administered an experimental treatment (X_E) or the control treatment (X_c), and posttested (O_2). An interim pre- and posttest (P_1 and P_2) was administered to the experimental treatment group. This is discussed in the appropriate section of this chapter.

A hazard encountered with the chosen design, suggests Borg and Gall (1971, p. 388) is the inability to attribute any observed differences noted by posttest results due to the lack of baseline data normally determined with a pretest measurement. For this reason, the subjects used were those who stated they had no previous experience using any automated information retrieval system.

This assumption is not detrimental to the investigation of the problem, since in the past two years almost all users of ERIC/CRESS computer search facilities have been novices. Thus all subjects, it is assumed, have the same level of knowledge of the ERIC/QUERY information retrieval system--very little, if any.

Setting of the Study

This experiment was conducted at ERIC/CRESS, located one hundred yards from the newly constructed building housing the College of Education on the NMSU campus. Collection of data began April 21, 1972, and continued through July 11, 1972. This time span allowed the selection of subjects attending long-term semester courses as well as summer courses.

Subjects assigned to the experimental treatment (EQUIP) were furnished a new IBM 2741 telecommunications terminal located in a well-lighted, adequately heated and air-conditioned, acoustically damped room. The only possible disturbance a subject might experience was the occasional noise produced by an IBM 029 keypunch machine operated in an adjoining room. Each subject occupied the terminal room by himself. Since a complete EQUIP session might take over two and one-half hours, each subject was allowed to set the time he was to use the terminal.

Subjects assigned to the control treatment were interviewed by an ERIC/CRESS information specialist in equally comfortable surroundings. Interviews were scheduled at the convenience of the subject.

Summary of the Experimental Design

Subjects desiring to use the ERIC/QUERY information retrieval system were divided into two groups--control and experimental. The experimental treatments are graphically portrayed in Figure 4.

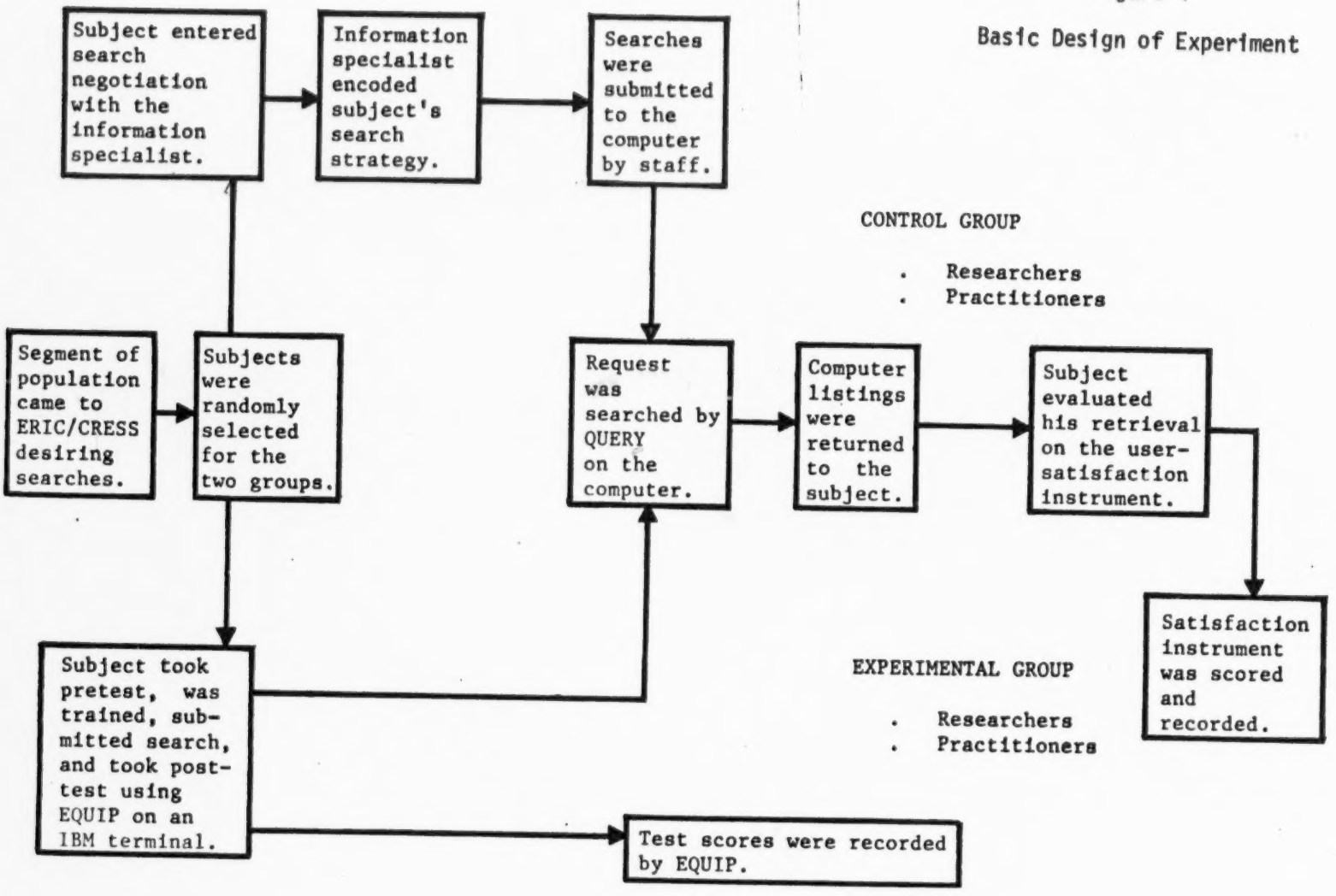
Both the experimental and control groups underwent a sequence of steps, shown as blocks in Figure 4, leading to the retrieval of a specific set of surrogated documents (citations) which were intended to meet individual informational needs. As depicted in Figure 4, the control group sequence (top activities) differed from that of the experimental group (bottom activities). Activities common to both groups are shown in the center of the paradigm. The activities which differ, search negotiations versus CAI training via EQUIP, were those of pivotal interest to the study.

Sampling Design

The study population. The population from which the study sample was drawn consisted of that group of people who have a need for information in the educational realm and are researchers, practitioners, or in training for the same. Additional population characteristics were: (a) male or female adults, (b) multi-ethnic, (c) must not have used ERIC/QUERY system previously, (d) must have been accommodated by the ERIC/QUERY facility, (e) must have been willing to spend the appropriate amount of time necessary to accomplish a search in person, and (f) must have been aware of ERIC/QUERY facilities. There is no reason to believe that the population described here would be different from persons seeking similar

Figure 4

Basic Design of Experiment


Subject entered search negotiation with the information specialist.
Information specialist encoded subject's search strategy.
Searches were submitted to the computer by staff.
CONTROL GROUP
. Researchers
. Practitioners
Segment of population came to ERIC/CRESS desiring searches.
Subjects were randomly selected for the two groups.
Request was searched by QUERY on the computer.
Computer listings were returned to the subject.
Subject evaluated his retrieval on the user-satisfaction instrument.
Satisfaction instrument was scored and recorded.
Subject took pretest, was trained, submitted search, and took post-test using EQUIP on an IBM terminal.
EXPERIMENTAL GROUP
. Researchers
. Practitioners
Test scores were recorded by EQUIP.

information in other geographic locations of the nation.

Sampling technique. Since it was not possible to predefine a segment of the treatment population in terms of specific individuals, a modification of the systematic sampling method suggested by Sax (1968, pp. 140-1) and described in subsequent paragraphs was used.

A total sample size of sixty, thirty in each treatment group, was deemed necessary. Prior to conducting the experiment, a sequential randomized sign-in sheet was created using the random number generator feature of APL. The steps of this procedure were:

(1) A randomly generated vector of sixty numbers ranging from 1 to 60 was created.

(2) Every even position number within the random sixty-number vector was isolated from odd position numbers, thus creating two separate random independent vectors of thirty numbers each.

(3) A coin was tossed to determine which thirty-number vector would represent the experimental group. In this manner, one vector of numbers was assigned to the control group and the second was assigned to the experimental group.

(4) After "tagging" the numbers in each of the two vectors with an S (search negotiation) or T (terminal) respectively, the sixty numbers were sorted and merged to create a randomized treatment assignment sign-in sheet.

Persons coming to ERIC/CRESS (strictly on a voluntary basis) who expressed their desire to use ERIC/QUERY were required to sign in on the sequential sign-in sheet described above, after it was

ascertained by ERIC/CRESS staff that they were members of the study population. In this manner each subject was assigned to a treatment group at random. Only chance determined the treatment assigned to him.

Limits of inferences to the study population. Although there is no reason to believe that the population described previously does not exist elsewhere, it would be hazardous to attempt to draw inferences from any population other than that which has access to ERIC/CRESS. The sampling design employed allows inferences to be made only of this local set of users. No provisions were made for including persons from other locations of the country in the sampling design; therefore, any generalizations reported are confined to the characteristics of the members of the educational sector of the local study area.

Measures Employed

Because of the nature of the search negotiations process (the control treatment) it was impossible to administer a pre- or posttest to members of the control group to determine any gain of knowledge of the ERIC/QUERY system. Since EQUIP was designed to impart knowledge gain, pre- and posttesting was confined to subjects in the experimental group. By this procedure, it was possible to determine if EQUIP was a successful CAI tool.

To determine if any difference in user satisfaction of search effort was attributable to the experimental treatment, a second needed measure was accomplished at the termination of the treatment through the use of a user satisfaction questionnaire,

which was created and administered to both treatment groups.

A description of the control group pre- and posttest will be discussed first. Similarly, the user satisfaction instrument will be detailed last.

Pre- and posttesting of EQUIP. Due to the uniqueness of EQUIP and the necessity of measuring produced knowledge gains, coupled with the pretest switching designed into the programming effort, one of the first considerations was that of creating an instrument which could discriminate between a population possessing ERIC/QUERY knowledge and skills and one which did not possess such attributes. The steps used to create the pre- and posttests (identical instruments) were:

(1) A "pool" of over thirty items was created based on the author's experience with the ERIC/QUERY system.

(2) A prototype pretest was created on the terminal using APL.

(3) Volunteers from ERIC/CRESS (members of the knowledgeable population) and other volunteers not associated with ERIC/CRESS (unknowledgeable population) were administered the prototype.

(4) Prototype items which did not discriminate between knowledgeable and unknowledgeable subjects were deleted and/or replaced with new items until eighteen items, each meeting the measurement criteria relating to the major design objectives of EQUIP, were found to discriminate in the desired manner.

(5) The final instrument was again administered to different members of the previous population and each item was found to discriminate; thus concurrent validity was assured.

It will be noted in the instructional sequence provided in Figure 1 (Chapter 3) that the pretest is administered to members of the experimental treatment group immediately after competency in the use of the terminal has been demonstrated by the subjects. Similarly, the posttest, which is identical to the pretest, is administered after the subject completes his training session on EQUIP.

Using Kuder-Richardson Formula 20, in the manner suggested by Sax (1968, p. 161), the final session of the pre- and posttest used with the experimental treatment group had a reliability of .73. It should be noted that the EQUIP pre- and posttest instruments are "power" and not "speed" tests. The completed instrument is presented in Appendix B as part of the session sample.

User satisfaction instrument. The experimental design employed by this study required that a comparative measure of treatment groups be administered. Again, due to the nature of the study problem, it was impossible to administer a pretest observation to both groups. If a subject had not been served there would be no value to measuring his satisfaction with the service, since no service would have been received. Thus there was a need for only a post-observational measurement to determine if the treatments used fostered any differences in satisfaction between treatment populations.

It was rationalized that when a user arrives at ERIC/CRESS, he has several expectations related to his specific informational need:

(1) He desires at least some hits as a result of his effort.

(2) He wants document citations which are relevant to his information need.

(3) He expects to expend time in his search effort.

(4) He expects ERIC/CRESS to assist in his search effort.

(5) He expects his total information research effort to be interesting, challenging, and useful.

The extent to which the user's information retrieval experience fostered positive and/or negative reactions within each of the criterial characteristics was measured with the simple Likert-type questionnaire provided in Appendix C. The questionnaire, the user satisfaction instrument discussed previously, was created and tested with a small sample of subjects who had just received the results of an ERIC/QUERY computer search. Concurrent and content validity were established via unstructured interviews with these subjects.

Collecting Additional Information

After each subject signed in (described previously), he was directed to complete a one-page form giving his name, subject major (if any), college affiliation, phone number, and academic classification. Additionally, he was requested to complete a statement of informational need and to indicate his retrieval expectations as related to number of hits and relevancy of his search. This information collection form may be found in Appendix C.

Chapter 5: Analysis of Data

The prime thesis of this study directly relates to three specific areas of inquiry: (a) Is it possible to construct a viable user-to-information-system interface in a CAI environment? (b) Will uninitiated users of such a CAI interface be as satisfied with information retrieval results as those undergoing traditional search negotiations with an intermediary information specialist? and (c) Will users expecting only a few highly relevant hits (practitioners) be as satisfied with their search effort as those who desire a large number of multi-relevant hits (researchers)?

The analysis of data collected for the study will be described in five major sections of this chapter. In the first section a detailed description of the sample which underwent the study treatments is presented. The second section delineates comparisons between experimental and control groups. In the third section the performance of the experimental group is discussed. Additional findings related to the scope of the study are given in the fourth section, and a brief summary is provided in the last section.

Description of the Sample

Types of users. Lancaster (1968, p. 59) and the Rand Corporation (1971, pp. 20-3) strongly advocate functional differences between the requesters of information from an information retrieval system. Generally, users of these systems are classified as researchers and practitioners. The distinction between these groups is the expectation each has of the system. Since the literature did

not provide any clear definitions for these terms, it was felt that a more succinct description of the sample should be extended in this chapter of the report.

As given in the operational definitions (Chapter 1), a researcher is defined as a user seeking a large number of relevant to slightly relevant hits, and a practitioner is a user seeking a small number of highly relevant hits as a result of search effort. The two types of users, therefore, are separated only by the number of expected or desired hits retrieved and the relevancy of the hits retrieved. A problem lies in the meaning of "large" and "small" numbers of hits and "highly relevant" to "some relevancy". To place more direct meaning on these ambiguities, provisions were made for subjects to respond to choices of relevancy and retrieval volume (see Appendix C) in the manner described in Chapter 4. After all data was collected, a comparison of responses to the type of relevancy and number of hits desired was conducted. The results are provided in Table 6. It will be noted that the number of subjects indicating their desire for "all document citations . . . regardless of how relevant . . ." chose values ranging from 1-10 through 50+ as the number of desired hits. The same is true for subjects indicating they wanted "all . . . highly relevant" citations to be retrieved by the system. It may be concluded that the key word "all" is not perceived differently by the study sample responding to either of the first two choices on the computer search input information form. By inspection of Table 6 there can be seen a clear tendency toward expecting a greater number of hits from the ERIC/QUERY system by

these subjects, the researchers. A definite shift of agreement between the number of hits desired and the type of users is evident.

Table 6

Frequencies of Study Sample Subjects in Relation to Retrieval Expectations

Number of Hits Desired	Type of User			
	Researcher		Practitioner	
	Desires All Hits Regardless of Degree of Relevancy	Desires All Highly Relevant Hits	Desires Only a Few Highly Relevant Hits	Desires Only a Few Hits Regardless of Degree of Relevancy
1-10	3 (14.29)	3 (9.38)	1 (20.0)	0 (0)
11-20	3 (14.29)	16 (50.0)	3 (60.0)	0 (0)
21-50	8 (38.1)	12 (37.5)	1 (20.0)	0 (0)
50+	7 (33.3)	1 (3.12)	0 (0)	0 (0)
Total	21	32	5	0

Note.--Values given in parentheses are percentages based on column totals.

The above trend does not manifest itself with subjects indicating a desire for only a few highly relevant hits. There is an obvious difference, since no experimental subjects indicated the need for over 50 hits while at the same time expecting a "few" highly relevant hits. It is also interesting to note that of the 58 sample subjects, none indicated they wanted "only a few" hits regardless of relevancy.

It is felt that the foregoing discussion legitimizes the use of the terms "researcher" and "practitioner" as previously defined. Researchers wanted an "in-depth" search with numerous document citations; a practitioner did not.

User characteristics. For the reasons discussed in Chapter 4 under sampling technique, until the collection of data was completed no specific description of the study sample could be provided. Only after the experiment was completed and subsequent descriptive analysis applied could the sample be more clearly identified. Therefore, it is felt, such discussion is appropriate in this chapter of the report.

As stated in Chapter 4 (procedures), 60 novice subjects desiring to use the services of the ERIC/QUERY information retrieval system were assigned to either the control group (search negotiations) or the experimental group (CAI training and search submission via EQUIP). Of the sample of 60, 58 completed the experiment. Subject number 17 left the campus without completing the search evaluation form and subject number 47 did not complete the CAI training due to time commitments. Both these subjects had been assigned to the experimental group, but for the reasons stated, both were deleted from the sample, thus reducing the total study sample to 30 control and 28 experimental treatment subjects. There was no evidence indicating that the loss of two subjects to the experiment was due to negative conotations caused by the experimental treatment.

There was a total of 31 men and 27 women in the sample; 16 men and 14 women in the control group, 15 men and 13 women in the

experimental group. Table 7 shows these and other categories expressed as frequencies.

Table 7

Study Sample by Treatments

Characteristic				Treatment	
Type	Expectations	Sex	Academic Class	Control	Experimental
Researcher	Desiring Retrieval of all highly relevant document citations	Men	Graduate	5	9
			Undergraduate	1	1
		Women	Graduate	3	1
			Undergraduate	0	1
	Desiring retrieval of all related document citations	Men	Graduate	9	5
			Undergraduate	0	0
		Women	Graduate	6	8
			Undergraduate	3	1
Practitioner	Desiring retrieval of a few highly relevant document citations	Men	Graduate	0	0
			Undergraduate	1	0
		Women	Graduate	0	1
			Undergraduate	2	1
	Desiring retrieval of a few document citations regardless of relevancy	Men	Graduate	0	0
			Undergraduate	0	0
		Women	Graduate	0	0
			Undergraduate	0	0
Totals				30	28

When it is remembered that subjects were assigned to the respective treatment groups purely by chance, the quantities appearing in Table 8 reveal much about the population characteristics. Since no descriptive data of ERIC/QUERY users was found in the literature, it

Table 8

Numbers and Percentage of Various Categories of Subjects Within Treatment Groups

Category	Treatment Groups				Totals	
	Control N_T = 30		Experimental N_T = 28			
	N	%	N	%	N	%
Researchers	27	90.0	26	92.9	53	91.4
Practitioners	3	10.0	2	7.1	5	8.62
Males	16	53.3	13	46.4	29	50.0
Females	14	46.	15	53.6	29	50.0
Graduates	26	86.	24	85.7	50	86.2
Undergraduates	4	13.3	4	14.3	8	13.8
Male Researchers	15	50.0	15	53.6	30	51.7
Male Practitioners	1	3.33	0	0	1	1.72
Female Researchers	12	40.0	11	39.3	23	39.7
Female Practitioners	2	6.67	2	7.1	4	6.89
Male Graduate Researchers	14	46.7	14	50.0	28	48.3
Male Undergraduate Researchers	1	3.33	1	3.57	2	3.45
Female Graduate Researchers	9	30.0	9	32.1	18	31.0
Female Undergraduate Researchers	3	10.0	2	7.14	5	8.62
Male Graduate Practitioners	1	3.33	0	0	1	1.72
Male Undergraduate Practitioners	0	0	0	0	0	0
Female Graduate Practitioners	2	6.67	1	3.57	3	5.17
Female Undergraduate Practitioners	0	0	1	3.57	1	1.72

Note.--Percentages are given for the respective treatment groups while total percentages are based on the total sample (n = 58).

was decided, based on similarity of the results of the two independent random subsamples, to present the full set of categories depicted on Table 8, though only the first two categories are considered in depth as part of the analysis presented.

The difference between category frequencies shown on Table 8 are minute, with the greatest percentage difference found between female and male across treatment groups and the smallest between the percentage of female practitioners--a range of .47 to 6.9 percentage points. This clearly shows that the two subsamples are almost identical in characteristics, a fact which supports the sampling technique employed for the experimental design.

Comparison of Experimental and Control Groups

Analysis of measures of user satisfaction. Four of the five experiment hypotheses involve the use of satisfaction scores. Each hypothesis will be restated, followed by appropriate data analysis.

Comparison of user satisfaction item means. Treatment group scores obtained from the five Likert-type scales of the computer retrieval evaluation form were compared using the t statistic in the manner suggested by Winer (1962, pp. 31-33) for comparing uncorrelated means of two groups with unequal n's. To test the assumption of homogeneity of variance for each measure, the appropriate F ratios were computed as suggested by Winer (1962, pp. 33-6). The critical value for the two-tailed t statistic when $\alpha = .05$ and df = 56 ($df = N_a + N_b - 2$) is $t_{.975}(56) = 2.04$, a value obtained from Winer's tables (1962, p. 641). The F ratio critical value is $F_{.95}(28,30) = 1.87$ from the Chemical Rubber Company (CRC) Tables (1968, p. 308).

As was noted in Table 9, the assumption of equal variance of group scores for item three lacks support. This necessitated additional analysis of this single measure. Winer (1962, p. 37) suggests that if both group sizes are approximately thirty or greater ($N_a = N_b = 30$) the $\underline{t}'$ statistic is an appropriate test of H_o: $\bar{x}_1 = \bar{x}_2$. Accordingly, the computations for testing differences between independent sample means with unequal variances were applied to data collected for item three of the user satisfaction questionnaire. A summary of these calculations is given in Figure 5.

Figure 5

Summary Calculations of $\underline{t}'$ Statistic for Item Three of the Satisfaction Questionnaire

	Treatments	
	Control	Experimental
N	30	28
Means	3.7	3.11
Variance	.286	.692
H_0: $\mu c_3 = \mu e_3$	$\underline{t}'_{obs}$ = 3.188135	
H_1: $\mu c_3 \neq \mu e_3$	$\underline{c}$ = .28735	
	$\underline{f}$ = 45.53	

The values $\underline{c}$ and $\underline{f}$ are based on Welch's (Winer; 1962, p. 37) derivation and approximation of the Student's $\underline{t}$ distribution through a correction to the degrees of freedom which are used to ascertain the critical $\underline{t}$ value. Thus the critical value of $\underline{t}$, using the

Table 9

Summary of Statistical Computations
Based on User Satisfaction Scores

Item	$\bar{x}_C$ N=30	$\bar{x}_E$ N=28	$\bar{x}_C - \bar{x}_E$	F Ratio	$\underline{t}$	Hypothesis Tested ($H_1: \mu_{e_i} \neq \mu_{e_i}$)
1. Satisfaction of number of hits	3.00	2.89	.017	$\frac{.988}{.966} = 1.023$	0.413	$\mu_{e_1} = \mu_{e_1}$
2. Satisfaction of overall relevancy of hits	3.00	2.75	.250	$\frac{.787}{.621} = 1.267$	1.136	$\mu_{e_2} = \mu_{e_2}$
3. Satisfaction of time spent	3.7	3.11	.593	$\frac{.692}{.286} = 2.417^a$	3.251^a	$\mu_{e_3} = \mu_{e_3}$
4. Satisfaction of clearing-house effort	3.67	3.82	-.155	$\frac{.300}{.299} = 1.003$	-1.076	$\mu_{e_4} = \mu_{e_4}$
5. General Satisfaction	3.13	3.00	.133	$\frac{.815}{.602} = 1.354$	0.604	$\mu_{e_5} = \mu_{e_5}$

[a]The F ratio exceeds the critical value of $F_{.95}(28,30) = 1.87$, Therefore $H_0: \sigma_{e_3}^2 = \sigma_{e_3}^2$ is rejected. This necessitated further analysis of the treatment means of this measure since the computed $\underline{t}$ might not be usable for inference purposes due to a homogeneity of variance violation.

appropriate conversion, becomes $t_{.975}(f) = t_{.975}(46) = 2.01$ for a two-tail test.

It can be seen that the computed t value was outside the $t_{.975}$ critical range. Thus it is concluded that there were differences manifested between the treatment populations as related to the amount of time spent obtaining search results. The testing of hypothesis one relating to treatment mean differences as measured by analysis of variance (AOV) using total satisfaction scores is provided in the next section.

Hypotheses One, Four, and Five

It was shown in the first section of the chapter that due to user expectations of an information retrieval system, they may be classified as researchers or practitioners. Hypotheses one, four, and five, restated below, were tested with the application of analysis of variance (AOV) technique to determine if a difference in interface experiences were related to treatment effect or user type.

(1) There are no significant differences between the measured user satisfaction with computer search results of CAI-interfaced subjects and those undergoing search negotiations with an information specialist.

(4) There are no significant differences in user's satisfaction between practitioners and researchers as determined by the user satisfaction instrument.

(5) There are no significant interactions between types of users (practitioner or researcher) and treatments (CAI or search negotiations) as measured by the user satisfaction instrument.

Prior to selecting an appropriate analysis technique, the following data table was created:

Table 10

Summary Table for Experimental and Control Treatment Groups Based on Total Satisfaction Scores Obtained on the User Satisfaction Instrument

Total N=58	Researchers				Practitioners	
	B_1 Subjects Desiring all citations regardless of relevancy		B_2 Subjects Desiring all highly relevant citations		B_3 Subjects Desiring a few highly relevant citations	
	N	$\bar{Y}$	N	$\bar{Y}$	N	$\bar{Y}$
a_1 (N=28) Experimental Group (EQUIP)	12	15.16667	14	15.4285	2	19.0
a_2 (N=30) Control Group	9	16.7778	18	16.6111	3	15.00

The design of the study can be viewed as having a 2 x 3 factorial set of treatments. It can be seen that the six cells of the summary table display unequal and disproportionate subclass membership; but in such a way that the analysis proposed by Myers

is defensable. Myers states:

> Disproportionate cell frequencies may be representative of treatment population frequencies, particularly when the independent variable is an individual characteristic such as age, sex, socio-economic status, a personality trait, or intelligence. In such cases, the method of expected cell frequencies may provide an appropriate analysis [1966, p. 190].

Since a user's expectations from the information retrieval system are deemed internal (such as personality and intelligence) and independent, the AOV methodology suggested by Myers (1966, p. 190) for disproportional cell frequencies with a 2 x 3 factorial experiment was followed exactly. This procedure differs from the more common least square solution in that cell frequencies and totals are recomputed based on the probability of sampling an experimental unit from a population of like experimental units. The results of the analysis are provided in Table 11.

UCA

Table 11

AOV for Testing Hypotheses One, Four, and Five

Source	df	SS	MS	F-value
Total	58	16371.000		
Mean	1	14990.07143		
A effect	1	12.096233	12.096233	.455
B effect	2	3.669421	1.8347105	.069
AB interaction	2	24.90709527	12.453547	.4689
Residual	52	1380.9287	26.5563	

Note.--The critical F values are $F_{.95}(2,52) = 3.15$, $F_{.95}(1,52) = 4.00$ which were not exceeded; thus the nonsignificance of the related effects.

Given the procedure and analysis described, it was concluded that there are no differences between the responses of the treatment groups (A effect), the types of users as classified by expected hits (B effect), or interaction between the two factors. Hypotheses one, four, and five may not be rejected.

Performance of the Experimental Group

Pre- and posttest measures. To provide evidence that the CAI training and search submission program (EQUIP) did increase the knowledge of the user, hypothesis two, re-stated below, was tested in the same manner as hypothesis one.

(2) There are no significant differences between the subjects' pre- and post-measures of knowledge of the information retrieval

system administered during the experimental CAI treatment.

Using the procedure suggested by Winer (1962, pp. 39-43) for testing hypothesis between two means with correlated observations which assume a linear additive model, calculations resulting in the values presented in Figure 6 were accomplished.

Figure 6

Summary of Calculations Comparing the Pre- and Posttest Interface Knowledge Gains of the Experimental Treatment Group

Hypothesis Tested	$\overline{X}_{post}$	$\overline{X}_{pre}$	$\overline{X}_{post}-\overline{X}_{pre}$	t_{obs}
$H_o: \mu_A = \mu_B$ $H_a: \mu_A \neq \mu_B$	79.786	34.071	45.714	14.09

Since $t_{.95}(28) = 2.05$, t_{obs}, $H_o: \mu_A = \mu_B$ is rejected and it is concluded that the difference of 45.714 points average gain from pre- and posttest measures could not have occurred by chance at $\alpha = .05$--hypothesis two is rejected.

Pre- and posttest scores related to satisfaction score.

Hypothesis three: There is no significant relationship between measured knowledge of the ERIC automated information retrieval system acquired as a result of CAI training and user satisfaction as measured by the user satisfaction instrument. To investigate the possibility of a simple linear regressive association between the pre- and posttest gains made by subjects of the experimental group and their respective total scores on the satisfaction questionnaire,

a test of hypothesis three was accomplished. If a predictive relationship could be found to exist, additional support for the treatment effect on user satisfaction would be gained. Using the APL linear regression program MREG provided by the NMSU computer center which assumes the linear model $\underline{Y} = \underline{XB} + \underline{E}$, two regression analysis procedures were accomplished.[1] A summary of these efforts is presented in Figure 7.

[1]The procedures for this analysis were those suggested by Draper and Smith (1966).

Figure 7

Summary of Linear Regression Computational Results Relating to the Predictive Relationship of Posttest Scores and Pre- and Posttest Gains to Experimental Subject's Satisfaction Scores

Relationship	Dependent Variable	Independent Variable	B_0 (Intercept)	B_1 (Slope)	F-value (Regression)	R^2 (%)
Total Pre- and posttest gains to total satisfaction score	Total Satisfaction Score	Gain Score	19.67498	-0.08977	6.38*	19.71
Posttest score to total satisfaction score	Total Satisfaction	Posttest Score	18.4557	-0.03615	0.45	1.69

Critical F value = $F_{.95}(1,25) = 4.23$

*Exceeds critical $F_{.95}$ value of 4.23

The results presented in Figure 7 agree with the scattergrams of the data. By inspecting the plotted values of pre- and posttest gains versus total satisfaction score, Figure 8, it is fairly obvious why the applied linear regression model accounts for only about 20 per cent of the total variation--very little if any pattern exists.

The previous discussion does not explain the significant regression manifested by the critical F value of the first calculation presented in Figure 7. By careful inspection of Figure 8 (scattergram of gains versus total satisfaction scores), two isolated clusters appear to exist, one in the top portion of the plot and a second in the bottom portion. It was felt that further investigation was warranted since the relationship of the total satisfaction scores to retrieval precision for the subjects isolated in the lower cluster of Figure 8 were highly correlated (See Figure 9).

It is interesting to note that all of the subjects contained in cluster 2 obtain a search precision of zero with the exception of subject number 18 who obtained a precision of 100. There is evidence that this subject's satisfaction was low because, even though he obtained maximal precision on his search, he received only nine hits but was expecting over 50. From this, we conclude that the cluster of subjects providing a lower total satisfaction score were those who became disenchanted with the experimental treatment because they received less than expected results from their information retrieval effort, thus influencing the linear regression analysis shown to be significant in Figure 7.

Figure 8

Scattergram of Experimental Group Total Satisfaction Scores Versus Pre- and Posttest Gains

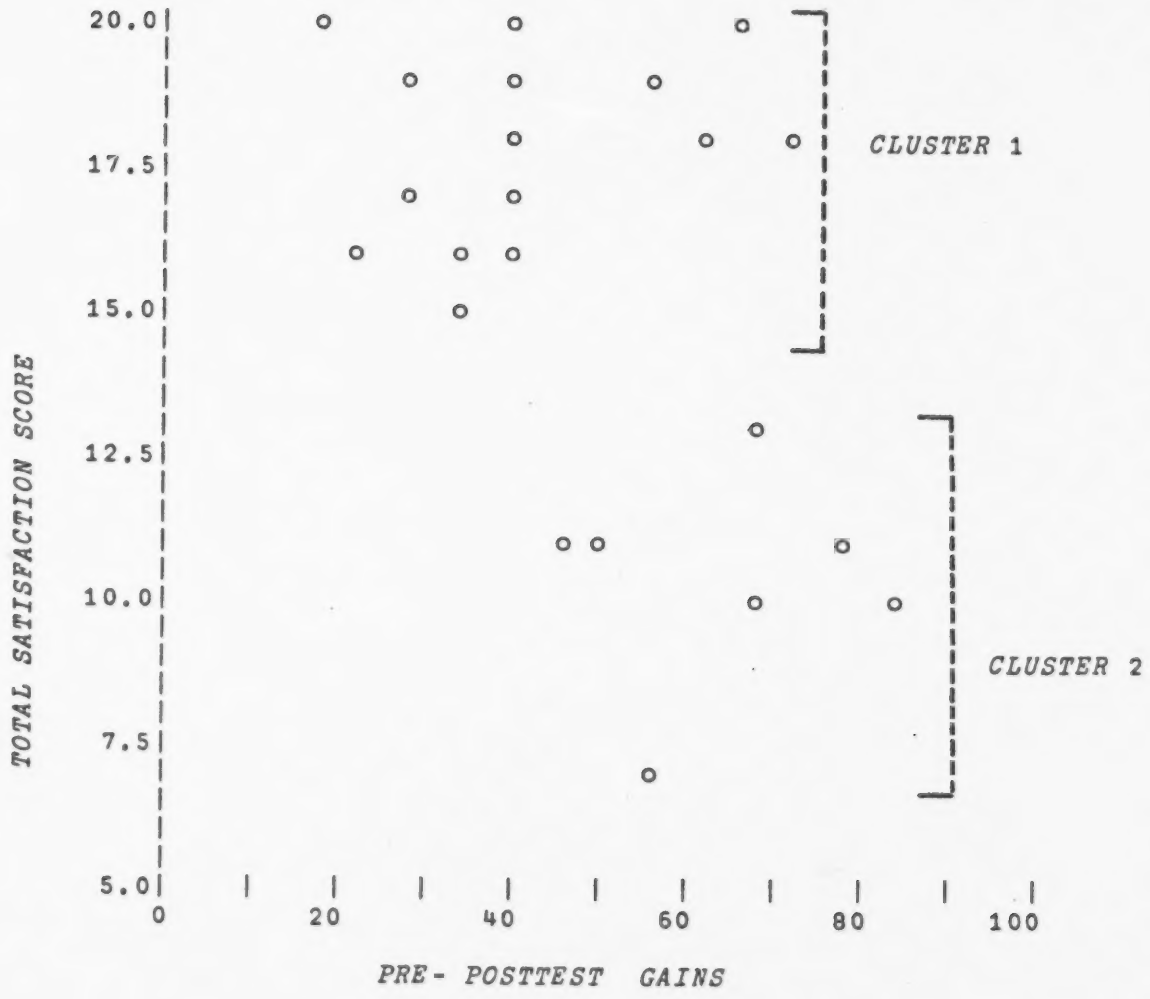


Figure 9

A Contrast and Correlation of Experimental Total Satisfaction Scores and Search Precision for Cluster Two

Subject Number	Total Satisfaction Score	Precision of Subjects' Search Effort
15	10	0
18	11	100
22	7	0
23	11	0
33	11	0
59	10	0
60	13	0

Correlation = r = .7095

Basically the question was asked Was the effect of this small cluster of subjects obtaining low precision for their search efforts responsible for a significant negative slope of the B_1 line of the regression model? To investigate this possibility further, the two groups of subjects in question were separated and the linear analysis technique previously described was applied to each to determine if any regression effects existed for the isolated cluster. The results of these computations are provided in Figure 10.

Figure 10

A Contrast of Linear Regression Summary Calculations for the Cluster 1 and Cluster 2 Subjects Shown on Figure 8

Cluster	Dependent Variable	Independent Variable	B_0 (Intercept)	B_1 (Slope)	F-value (Regression)
1 (Top) N=21	Total Satisfaction Score	Gain Score	15.8401	0.364	1.75801
2 (Bottom) N=7	Total Satisfaction Score	Gain	9.636	0.0124	.0460

Critical F value for cluster 1 = $F_{.95}(1,21) = 4.30$

Critical F value for cluster 2 = $F_{.95}(1,7) = 5.59$

Supported by the analysis, it is concluded that the existence of a second different sample group within the experimental sample (cluster 2, Figure 8) is the cause of a significant regression effect at $\alpha = .05$ when the entire experimental sample's dependent total satisfaction scores are predicted using a linear model. However, because the R^2 value of 19.7 indicates that the model accounts for only approximately 20 per cent of the variability of the total sample responses, and because two clusters, both lacking any significant regression effect were shown to exist, hypothesis three may not be rejected.

Additional Findings

In addition to the observational data applied to each of the experimental hypotheses previously discussed, careful records of

the retrieval precision of each subject were maintained. The precision of a search effort is expressed as a ratio converted to a percentage of the number of retrieved hits to the total number of citations retrieved by the ERIC/QUERY system. For example, a precision value of 50 is interpreted as "50 per cent of the citations retrieved were considered by the user to be relevant to his given information need." Both Lancaster (1968) and Caruso (1969) refer to the precision ratio as one measure of an information retrieval system desired traits; i.e., the better the ratio the greater utility of the system.

Considerable experience with the ERIC/QUERY system by the author has lead to the conclusion that the precision of the ERIC/QUERY system is less than ideal. Part of the lack of efficacy is attributed to the massive size of the thesaurus of descriptors used with the system. There are over 4,000 ERIC descriptors which, when modified via the TEXT function of ERIC/QUERY may be expanded almost to infinity.

It was felt that analysis of user's precision values was worthwhile. More specifically, does the retrieval precision of an expert (the information specialist) differ from that of a novice (the experimental treatment sample)? More formally, this question may be stated as a statistical hypothesis: There are no differences between the average precision rates of users undergoing CAI training and submission via EQUIP and those obtaining the expert services of an information specialist.

To test this hypothesis, Student's $\underline{t}$ was applied, giving the results presented below.

Figure 11

Summary of Statistical Computation for Determinining the Significance of Experimental Group Precision Means

X_c	X_E	$X_c - X_E$	$\underline{t}$	F-ratio
58.2	52.679	5.521	.554	.667

Note.--The F_{obs} ratio does not exceed $F_{.95}(28,30) = 1.87$ and $\underline{t}_{obs}$ does not exceed $t_{.975}(56) = 2.04$.

The methodology used to obtain the values presented in Figure 11 was identical to that described for testing hypothesis one. Based on the non-significant value of $\underline{t}_{obs}$, it is concluded that $H_o: \mu_c = \mu_e$ based on observed precision values may not be rejected. It makes no difference whether a search is encoded by an expert or a novice trained on a CAI interface program; ERIC/QUERY retrieval precision is uniformly low--about 55 per cent.

Summary

A complete description of the experimental sample was presented. Each of five major experimental hypotheses were tested. Of the five, only number two was rejected. A sixth hypothesis relating to search precision was also tested. Additionally, the use of the terms researchers and practitioners as applied to users was legitimized.

Chapter 6 of this report will present an overall summary of the experiment, conclusions, recommendations for further study, and the specific implications of the statistical tests already presented.

Chapter 6: Summary, Conclusions, and Recommendations

Summary

The prime objectives of the study, reduced to more elemental partitions, were as follows:

(1) To develop, implement, and test a fully documented CAI package for training potential users of ERIC/QUERY to interface with the existing system.

(2) To test the effectiveness of the CAI program produced (EQUIP) with the use of pre- and posttest measures.

(3) To provide an instrument for measuring user satisfaction of his search effort.

(4) To determine if interface methods (search negotiations--CAI training and search submission) cause a significant difference in user satisfaction of experimental and control groups.

(5) To determine if any interaction between treatment and user-type (researcher and practitioner) factors exist.

Each of the objectives was accomplished resulting in the interpretations offered in the second section and the conclusions discussed in the third section of this chapter. The fourth section explores the possible implication of study results and the last section delineates the author's recommendations.

Interpretations of Experimental Results

Among treatments. To answer the pivotal issue of whether or not a specially constructed CAI interface program is viable, statistical tests were conducted to determine if population

differences measured by the users' satisfaction responses existed. Each of these tests was described in Chapter 4 as related to hypotheses one, four, and five.

Statistical results of the testing procedure clearly indicated that the two interface methodologies did not cause any measurable differences in user's total satisfaction with the results of his search efforts. Analysis of individual items of the satisfaction questionnaire showed that, of the five, only one relating to the amount of time spent obtaining a search measured any differences between the treatment populations. The experimental group was slightly less satisfied with the approximately two and one-half hours of time spent using the terminal as compared with control group members which spent less than thirty minutes in the search negotiation interview.

It is felt that the measurable difference in satisfaction of the amount of time spent by subjects of both treatment groups is non-significant in a practical sense. It will be noted in Table 9 on page 50 that both mean satisfaction scores for item three, $\overline{x}_C = 3.7$ and $\overline{x}_E = 3.11$, are above the mid-point (2.5) of the Likert-type scale, suggesting that neither experiment group was dissatisfied with the time spent obtaining a search.

Additionally, experimental group subjects returning to use EQUIP for a second or third time (on their own and apart from the study) were observed initiating searches via a non-training version of EQUIP in less than 15 minutes. It is assumed that these subjects gained enough skill as a result of their initial usage of EQUIP to

accomplish a viable search. This phenomenon has strong implications for future research relating to the time a user will spend obtaining a search on EQUIP after he has been initially trained.

It should be noted that EQUIP, as used for this study, consumes almost 25 minutes of the user's time (after he has submitted his search) for posttest purposes. Unless used for a controlled study, the posttest should not be necessary in future versions of EQUIP.

No interaction between types of users, researchers or practitioners, was found to exist. This shows that both types of users were equally satisfied with their search efforts and results regardless of interface methodology (treatment).

Within the experimental treatment. A test of the third hypothesis relating to pre- and posttest gains of CAI interface users indicates that EQUIP does cause a significant increase in the novice user's ERIC/QUERY computer-searching skills. Surprisingly, however, there is no predictive relationship between such gains and the precision of the resulting searches when tested through linear regression techniques (hypothesis three, Chapter 4). Such a lack of predictability, it was felt, was due to the fact that precision attributes of ERIC/QUERY were such that, regardless of the computer-searching skill needed to execute a viable search strategy, precision rates remained less than ideal. To test this possibility, a separate analysis, reported in the latter part of Chapter 5, was conducted to determine if the search precision of assumed experts was equal to that of novice users trained on the telecommunication terminal.

Interestingly, no difference between mean search precision of the two groups could be detected. This phenomenon may be due to any of three possibilities: (a) the novice users were as skilled as the information specialist; (b) regardless of skill, ERIC/QUERY is uniformly inefficient; or (c) a combination of both. Unfortunately, no empirical evidence may be brought forth which rectifies this issue; however, the results of the precision test between groups supports the use of EQUIP.

In addition to conducting linear regression analysis to determine possible predictive relationships between EQUIP gain scores and user satisfaction, separate regression analysis was conducted for two cluster groups shown to exist within the sample--those subjects with very low retrieval rates (cluster 2) and all others (cluster 1). Due to the existence of the second cluster, it is evident that total satisfaction scores, and thus the significance of total subsample regression line, is effected, since separate regression analysis for each cluster was determined to be non-significant. This clearly shows the effect of some users' resulting search precision on their overall satisfaction of search effort--an understandable situation since users do want access to information.

Conclusions

Based on the procedures and analyses of collected data within the scope of the study, the following specific conclusions were formulated. These generalizations are confined to the specific population used in the study.

The CAI interface. It is not only possible to use a CAI-medium teleprocessing interface with which to bring novice users into direct contact with a complex computerized information retrieval system, but such an interface method (EQUIP) was created and demonstrated to function just as well as the search negotiations method in use at many installations. It was also demonstrated that populations of users classified as researchers or practitioners, based on pre-search retrieval expectations of the system, are both accommodated with the CAI interface in terms of their satisfaction with literature-searching efforts. Thus, the CAI interface services both types of users equally well.

The effectiveness of EQUIP. Although EQUIP is one of many possible approaches to training users to submit a computer literature search, the significant pre- and posttest means coupled with the fact that all of the subjects in the experimental group were able to submit a search which was not rejected by the system clearly demonstrates the viability of such an interface method. It is concluded that EQUIP, as an interface process, did meet the objective for which it was created.

User characteristics. Through the environment of a controlled experiment it is evident that at least two types of users, researchers and practitioners as defined in this study, may be expected to take advantage of automated information retrieval capabilities. It is also concluded that the retrieval expectations of these two classes of users are different.

General. Tests of each of four main study hypotheses designed to examine relationship between the experimental and treatment population resulted in the inability to reject the null hypotheses; i.e., no differences were found between the experiment groups within the dimensions studied. It is concluded that the CAI interface method, when used in conjunction with ERIC/QUERY, was no more nor less effective than search negotiations interface even though knowledge gains were shown to exist among EQUIP users.

Implications.

The use of EQUIP. For the past two years the user desiring ERIC/QUERY retrieval services had as his only alternative for initiating a search, the services of an information specialist--a situation which is manifested in most computerized information retrieval systems desiring to continue receiving the advantages of batch processing. With the application of a CAI training and search submission interface, new alternatives will be created. The time utilized by an information specialist to initiate searches for users may confidently be relegated to the student and the computer. The number of users accommodated may be increased in direct proportion to the number of accessible telecommunication terminals available; thus a larger number of information seekers may gain access to information retrieval files at a time when the information need is critical.

It should be noted that the use of EQUIP, or similar CAI training and submission programs, is not limited to the ERIC/QUERY system. EQUIP, with the necessary modifications, may be converted

for use with any batch processing information retrieval system if minimum hardware and software requirements are provided--specifically, similar APL capability used for the study. Whether or not a modified EQUIP or similar CAI interface will be effective with other information retrieval systems, of course, is a matter for further study.

Utilization of user characteristics. The detailed sample description provided in the first section of Chapter 5 should be of some value to future planning related to ERIC/CRESS services. Additionally, the application of known user requirements coupled with the characteristics provided as a result of the study should be of substantial assistance in forecasting user requirements of computer facilities. Finally, user reaction to CAI training, as shown by their satisfaction scores, were positive--an encouragement to explore other CAI/APL applications at New Mexico State University.

Recommendations

Before the results of the study may be generalized for other areas of the nation, it is recommended that other institutions possessing the required capabilities replicate the study. As was initially stated in Chapter 1, there is no reason to believe that user population in other geographic areas differs from those from which the sample was drawn; however, the possibility of detectable population differences does exist.

From the view of one not unfamiliar with APL programming, EQUIP (Version I) as used in the study should be re-encoded. Version I consumes approximately twice the necessary storage

requirements. It is recommended that each function (subprogram) be "compacted" via the use of more efficient program logic and APL capabilities. The posttest portion of EQUIP, which consumes almost 20,000 bytes of storage and about 25 minutes of user's time, should be deleted if further testing of the system is not anticipated.

Absolutely no provisions were made for contrasting cost differences between the two treatments of the experiment--a highly recommended priority for future study.

Finally, it is strongly recommended that EQUIP or similar interface packages be made available to all members of the education sector on an "active" basis in contrast to the traditional "passive" approach observed in most library settings. Literally millions of dollars have been spent creating large, complex, automated information systems. Through the use of a viable CAI interface, the decision-maker, researcher, practitioner, student, or teacher would have no one to blame for not accessing needed information but himself.

References

Borg, W. R., & Gall, M. D. Educational Research, An Introduction. New York: David McKay Company, 1971.

Brandhorst, W. T., & Marra, S. J. Characteristics of the ERIC Data Base. Material for Distribution at the ERIC Data Base APL QUERY User Seminar, Raleigh, North Carolina, November 30-December 2, 1970.

Burchinal, L. G. The Educational Resources Information Center: An Emergent National System. Journal of Educational Data Processing, 1970, 7(2).

Caruso, D. E. F. "An Experiment to Determine the Effectiveness of an Interactive Tutorial Program, Implemented on the Time Sharing IBM System 360, Model 50, in Teaching a Subject-orientated User to Formulate Inquiry Statements to a Computerized On-line Information Retrieval System." Unpublished doctoral dissertation, University of Pittsburgh, 1969.

Chemical Rubber Company. Handbook of Tables for Probability and Statistics. (2nd ed.) Cleveland, Ohio: Author, 1968.

Computer Resources Corporation. QUERY, A Universal Search System--Users Manual. (3rd ed.) McLean, Virginia: Author, 1970.

Draper, N. R. & Smith, H. Applied Regression Analysis. New York: John Wiley and Sons, Inc., 1966.

Griffiths, D. G. Administrative Theory. New York: Appleton Century, 1959.

Halpin, A. W. Theory and Research in Administration. New York: McMillan Co., 1966.

Janda, K. Information Retrieval; Applications to Political Science. Indianapolis, Indiana: Bobbs-Merrill, 1968.

Kepner, C. H. & Tregoe, B. B. The Rational Manager. New York: Houghton Co., 1953.

Lancaster, F. W. Information Retrieval Systems. New York: John Wiley & Sons, Inc., 1968.

Lockheed Company. Users Manual, ERIC/DIALOG On-line Retrieval System. Sunnyvale, California: Lockheed Aircraft Corporation, 1971. (a)

Lockheed Information Sciences. ERIC/DIALOG. (Brochure) Palo Alto, California: Author, 1971. (b)

Meadow, C. T. The Analysis of Information Systems, A Programmer's Introduction to Information Retrieval. New York: John Wiley & Sons, Inc., 1967.

Meyer, P. Your Personal Success Planner, Volume I. Waco, Texas: Success Motivation Institute, Inc., 1968.

Myers, J. L. Fundamentals of Experimental Design. Boston: Allyn and Bacon, 1966.

Optner, S. L. Systems Analysis for Business Management. Englewood Cliffs, New Jersey: Prentice Hall, 1960.

Rand Corporation. "Alternative Models for ERIC Clearinghouse Network." Unpublished draft report. Santa Monica, California: Author, 1971.

Rubinoff, M. (Ed.) Toward A National Information System. Washington, D. C.: Spartan Books, 1965.

Salton, G. Automatic Information Organization and Retrieval. New York: McGraw Hill, 1968.

Sax, G. Empirical Foundations of Educational Research. Englewood Cliffs, New Jersey: Prentice Hall, 1968.

Shannon, C. E. & Weaver, W. The Mathematical Theory of Communication. Urbana, Illinois: University of Illinois Press, 1959.

Stufflebeam, D. L. (Ed.) Educational Evaluation & Decision Making. Itasca, Illinois: F. E. Peacock, 1971.

Sutter, E. M. G. "Individual Differences and Social Conditions as They Effect Learning Computer-Assisted Instruction." Unpublished doctoral dissertation, University of Texas at Austin, 1967.

Timbie, M. & Coombs, D. H. An Interactive Information System; Case Studies on the Use of DIALOG to Search the ERIC Document File. Stanford, California: ERIC Clearinghouse on Educational Media and Technology, 1969, ED 034 431.

Williams, F. W. Principles of Automated Information Retrieval. Elmherst, Illinois: The Business Press, 1965.

Winer, B. J. Statistical Principles in Experimental Design. New York: McGraw-Hill, 1962.

Appendix A

EQUIP

DELETED

For Further Information Contact:

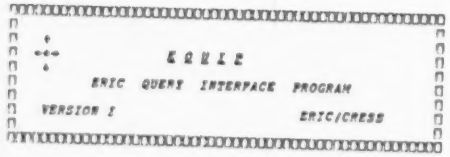
Dr. Albert D. Sink
State Department of Education
Room 120F
Rutledge Office Building
Columbia, South Carolina 29201

79/166

cat

Appendix B

Sample Session



1:41 PM WEDNESDAY JUNE 21, 1972

HI ! YOU HAVE JUST SIGNED ON TO A TERMINAL TRAINING PACKAGE DESIGNED TO BE OF ASSISTANCE
TO YOU IN YOUR SEARCH FOR RELEVANT INFORMATION FROM THE ERIC FILES. BEFORE YOU START, CONSIDER THIS - THE
TIME YOU SPEND NOW LEARNING HOW TO RETRIEVE INFORMATION PROPERLY FROM THE ERIC FILE IS BUT A FRACTION
OF THE TIME THAT IT WOULD TAKE YOU TO DO A REALLY GOOD SEARCH BY HAND! IN THE INTEREST
OF SAVING YOU TIME AND IMPROVING THE RESULTS OF YOUR SEARCH EFFORTS, THIS TRAINING PROGRAM HAS BEEN PREPARED.

DO YOU KNOW HOW TO USE THIS TERMINAL IN THE APL MODE (HOW TO CORRECT MISTAKES, ETC.)? (ANSWER BY TYPING THE
WORD YES OR NO AND HITTING THE RETURN BUTTON (ON THE RIGHT SIDE OF THE KEYBOARD)).

NO

OK. THE TERMINAL IS JUST AN ELECTRIC TYPEWRITER COUPLED TO THE COMPUTER.
IF YOU LOOK AT THE KEYBOARD, YOU WILL FIND THAT IT IS NOT TOO DIFFERENT FROM THAT OF
TYPEWRITER.

YOU HAVE ALREADY LEARNED THAT TO ENTER WORDS ON THE TERMINAL, YOU TYPE THE WORD AND THEN HIT THE
RETURN BUTTON. THIS SAME SEQUENCE TAKES PLACE FOR NUMBERS ALSO.

THERE ARE NO UPPER AND LOWER CASE LETTERS. ALL LETTERS ARE UPPER CASE.

WHAT IF YOU MAKE A MISTAKE WHILE TYPING? ONE NICE THING ABOUT A TERMINAL IS
THE EASE OF CORRECTING SUCH MISTAKES. THE PROCEDURE IS AS FOLLOWS:

STEP 1 - SPACE TO THE POSITION WHERE THE MISTAKE OCCURRED.
STEP 2 - HIT THE 'ATTENTION' [ATTN] BUTTON TO THE FAR RIGHT HAND SIDE OF THE KEYBOARD.
STEP 3 - AFTER THE TERMINAL HAS SPACED THE PAPER UP ONE TIME AND PRINTED THE SYMBOL '∨' DIRECTLY
UNDER THE SPACE WHERE THE MISTAKE WAS MADE, START RETYPING FROM THAT POINT.

AND THAT IS ALL THERE IS TO IT!

NOW TRY ONE. CORRECT THE SENTENCE BELOW.

THE USE OF THE TERMINAL IS SEMPLE.

```
                              ∨
                              MPLE
```

PLEASE TRY AGAIN, DON'T FORGET THE PERIOD.
THE USE OF THE TERMINAL IS SEMPLE.

```
                              ∨
                              MPLE.
```

SEE...IT IS NOT REALLY DIFFICULT AT ALL.

••• NOTICE ••• MANY MISTAKES ARE MADE BECAUSE MOST PEOPLE USE
THE LETTERS AND SYMBOLS PRINTED ON THE TOP OF THE KEYS. SINCE YOU SHOULD
BE USING ONLY APL SYMBOLS AND LETTERS, ALWAYS USE THE FRONT OF THE KEYS
AS YOUR REFERENCE...NOT THE TOP OF THE KEYS!

HAVE YOU EVER USED APL SYMBOLS (THE FUNNY LOOKING ONES ON THE FRONT OF THE KEYS) BEFORE?

NO
WELL, DON'T LET THIS BOTHER YOU TOO MUCH! VERY FEW SYMBOLS WILL BE USED.

YOU SHOULD FIND A KEYBOARD CHART BEHIND THIS TERMINAL.

IT IS SUGGESTED THAT YOU BECOME FAMILIAR WITH THE FOLLOWING SYMBOLS AND THEIR LOCATIONS (APPROXIMATE)
ON THE KEYBOARD.

SYMBOL	KEYBOARD LOCATION
(	JUST RIGHT OF THE LETTER L AS AN UPPER CASE.
)	JUST RIGHT OF THE SYMBOL (AS AN UPPER CASE.
⍬	JUST RIGHT OF THE 0 (ZERO) AS A LOWER CASE.
/	TO THE LOWER LEFT JUST LEFT OF THE SHIFT KEY AS A LOWER CASE.
*	ON THE 'P' KEY AS AN UPPER CASE.
⍳	JUST RIGHT OF THE LETTER W AS AN UPPER CASE.
⍴	JUST RIGHT OF THE LETTER W AS A LOWER CASE.

DON'T WORRY ABOUT ALL OF THE OTHER WEIRD SYMBOLS ... WE WILL NOT USE THEM!

NOW, TYPE THIS SERIES OF SYMBOLS AS SHOWN.(DO NOT USE SPACES,PLEASE.)

```
()∘/⍳1
()∘/⍳1
```

GOOD. JUST REMEMBER THAT THERE ARE NO UPPER AND LOWER CASE LETTERS. ALL LETTERS ARE UPPER CASE AND NO SHIFTING IS REQUIRED.

••• NOTICE ••• ALWAYS USE APL SYMBOLS AND LETTERS. DO NOT USE THE TOP OF THE KEYS AS YOUR REFERENCE! ALWAYS USE THE FRONT OF THE KEYS INSTEAD.

BEFORE WE GET STARTED INTO ACTUAL TRAINING, LET'S BRIEFLY GO OVER THE BEST WAY TO RESPOND TO THE VARIOUS ACTIVITIES OF THE TRAINING SESSION. ARE YOU READY? (TYPE YES OR NO)

YES

OK. ALL OF YOUR INPUT (VIA THE TERMINAL) WILL BE IN THE FORM OF SHORT WORDS, SHORT PHRASES, NUMBERS, SYMBOLS, SPECIAL LINES, AND/OR COMBINATIONS OF THESE. THERE WILL BE TIMES WHEN YOU ARE ASKED TO TYPE SPECIFIC RESPONSES. IN MOST CASES, THE DESIRED TYPE OF INPUT WILL BE CLEARLY ASKED FOR. IF A QUESTION IS ASKED WHICH REQUIRES A YES OR NO ANSWER, TYPE IN THE APPROPRIATE WORD. IN SITUATIONS WHERE A SELECTION IS TO BE MADE (MULTIPLE CHOICE), TYPE THE LETTER OR NUMBER FROM THE ALTERNATIVES GIVEN. IN CASES WHERE EXPLICIT SYMBOL GROUPS ARE ASKED FOR, COPY THEM EXACTLY!

SOME EXAMPLES FOLLOW:

QUESTION: DO YOU LIKE MUSIC?

YES

OK.
SITUATION: WHICH DO YOU THINK IS BETTER?
  [1] MONEY
  [2] A PLEASANT JOB
  [3] TIME TO PLAY

?

OK.

TO SAVE TIME, YOU WILL BE ASKED TO ANSWER A SERIES OF QUESTIONS WHICH WILL BE
USED TO CREATE A SPECIAL TRAINING SEQUENCE JUST FOR YOU. THERE ARE NO
''WRONG'' ANSWERS BECAUSE THE PURPOSE FOR THE QUESTIONS IS NOT TO OBTAIN JUST AN
ACHIEVEMENT SCORE. THE PURPOSE IS TO SET UP AN INDIVIDUAL PROGRAM FOR YOU.

PLEASE ANSWER THE FOLLOWING QUESTIONS HONESTLY. IF YOU DON'T KNOW AN ANSWER,
HIT THE RETURN BUTTON...DON'T GUESS!

<<< READ THESE INSTRUCTIONS CAREFULLY >>>
A TRE FOLLOWING QUESTIONS ARE OF 2 TYPES, TRUE OR FALSE , AND MULTIPLE CHOICE. ANSWER THE
A TRUE OR FALSE QUESTIONS WITH THE THE WORDS TRUE, OR FALSE. (DO NOT USE 'YES' OR 'NO' FOR THESE QUESTIONS.)
A ANSWER THE MULTIPLE CHOICE QUESTIONS BY SELECTING THE APPROPRIATE LETTER OR NUMBER.
A IN ALL CASES, IF YOU DON'T KNOW AN ANSWER, JUST HIT THE RETURN BUTTON!

CITATIONS OF DOCUMENTS ANNOUNCED IN BOTH RESEARCH IN EDUCATION (RIE) AND
CURRENT INDEX TO JOURNALS IN EDUCATION (CIJE) CAN BE RETRIEVED BY THE COMPUTER?

TRUE
SEARCHING FOR INFORMATION BY COMPUTER IS MORE EXACT (MUST BE BETTER PLANNED)
THAN SEARCHING IN A LIBRARY?

TRUE
TO OBTAIN RELEVANT INFORMATION FROM THE COMPUTER, YOU MUST KNOW
THE AUTHOR'S NAME OR THE TITLE OF THE DOCUMENT
DESIRED?

FALSE
A COMPUTER RETRIEVED DOCUMENT THAT YOU FEEL IS VERY
RELEVANT TO YOUR INFORMATION NEED IS CALLED A ''HIT''.
TO OBTAIN A SMALL NUMBER OF HITS FROM A SEARCH STRATEGY, IT IS
BETTER TO DEFINE YOUR SEARCH IN BROAD TERMS (THAT IS, USE
DESCRIPTORS WHICH FIT A LARGER CATEGORY OF DOCUMENTS)?

TRUE
WHICH OF THE FOLLOWING DESCRIPTORS WOULD BE MOST
APPROPRIATE FOR SEARCHING THE FILE FOR DOCUMENTS RELATING TO
''THE TEACHING OF DISADVANTAGED FIRST GRADE MEXICAN AMERICANS''?
(1) BILINGUAL EDUCATION (AND) GRADE 1
(2) BILINGUAL (AND) GRADE 1
(3) DISADVANTAGED (AND) MEXICAN AMERICANS (AND) TEACHING (AND) GRADE 1
(4) BILINGUAL EDUCATION (AND) GRADE 1 (AND) MEXICAN AMERICANS
(5) DISADVANTAGED
(6) MEXICAN AMERICANS (AND) TEACHING

3
WHICH OF THE FOLLOWING IS NOT A NECESSARY TOOL FOR FORMULATING A COMPUTER SEARCH?
(1) THE ROTATED THESAURUS
(2) THE THESAURUS OF ERIC DESCRIPTORS (''THE THESAURUS'')
(3) RESEARCH IN EDUCATION (RIE)
(4) THE TERM USAGE POSTINGS (''THE STATISTICS REPORT'')

4
(+) AND (/) ARE THE PROPER SYMBOLS FOR THE LOGIC
OPERATORS 'AND' AND 'OR'?

IN THE THESAURUS OF ERIC DESCRIPTORS, THE CODES ''RT'',''SN'',''BT'',
''UF'',''NT'', AND ''USE'' MAY BE FOUND UNDER THE DESCRIPTORS. WHICH
OF THE FOLLOWING IS AN INCORRECT DESCRIPTION FOR THE CODE GIVEN?
(1) RT = RELATED TERM
(2) SN = SINGLE NOUN
(3) BT = BROADER TERM
(4) UF = USE FOR
(5) NT = NARROWER TERM
(6) USE = USE THIS TERM

6
WHAT PART OF THE DOCUMENT RECORD (ON THE COMPUTER FILE) IS NORMALLY SEARCHED?
(1) THE AUTHOR'S NAME SECTION
(2) THE TITLE SECTION
(3) THE DESCRIPTOR SECTION
(4) THE ACCESSION NUMBER (''ED NUMBER'' OR ''EJ NUMBER'')

3
IF YOU SUBMIT A MECHANICALLY INCORRECT SEARCH,
(A) THE COMPUTER WILL BLOW UP
(B) YOUR SEARCH WILL RUN BUT WILL BE INVALID
(C) YOUR SEARCH WILL NOT RUN AND NO LISTINGS (RETRIEVED DOCUMENTS) WILL BE PRODUCED
(D) YOUR SEARCH WILL RUN BUT WILL TAKE TOO MUCH COMPUTER TIME
(E) THERE IS NO WAY OF KNOWING WHAT WILL HAPPEN

B
DESCRIPTORS MUST BE ENTERED ON YOUR SUBMITTED SEARCH
EXACTLY AS LISTED IN THE THESAURUS?

IN COMPUTER SEARCHING, A DESCRIPTOR NOT PRECEDED BY A STAR (*) WILL
RETRIEVE ALL DOCUMENT CITATIONS CONTAINING THAT PARTICULAR DESCRIPTOR --
REGARDLESS OF WHETHER MAJOR OR MINOR STATUS WAS ASSIGNED TO THAT DESCRIPTOR?

TRUE
THE COMPUTER WILL RETRIEVE 100 PERCENT OF THE DOCUMENTS
CONTAINING A PARTICULAR DESCRIPTOR?

THE FOLLOWING ARE SOME SEARCH STRATEGIES TO BE USED IN SUBSEQUENT QUESTIONS (ASSUME CORRECT CARD COLUMNS).

(A) 9999((TALLY)),J(*)CATS(TXT),(*)DOGS(TXT),(*)MICE(TXT). 01
(B) 9999((TALLY)),J(*)CATS(TXT),(/)DOGS(TXT),(/)MICE. 01
(C) 9999((TALLY)),J(*)CATS(TXT),(*)DOGS. 01
(D) 9999((TALLY)),J(*)CATS(TXT) 01

OF THE FOUR ENCODED SEARCH STRATEGIES GIVEN ABOVE,
WHICH IS CORRECTLY ENCODED? (A, B, C, OR D)

WHICH, IF IT WERE PROPERLY ENCODED, WOULD RETRIEVE THE LARGEST NUMBER OF HITS?
(A, B, C, OR D)

A
WHICH, IF IT WERE PROPERLY ENCODED, WOULD RETRIEVE THE FEWEST NUMBER OF HITS?
(A, B, C, OR D)

P
TO RETRIEVE THE DESIRED RESULTS FROM A COMPUTER SEARCH,
WHICH IS (ARE) THE MOST IMPORTANT CONSIDERATION(S)?
(A) KEEPING THE CHARACTERS IN PROPER CARD COLUMNS.
(B) SELECTING DESCRIPTORS.
(C) SELECTING THE LOGIC ORDER.
(D) ALL OF THE ABOVE.
(E) NONE OF THE ABOVE.

P
IF IN DOUBT ABOUT HOW A DESCRIPTOR HAS BEEN DEFINED, IT WOULD BE TO YOUR ADVANTAGE TO REFER TO:
(A) THE ROTATED THESAURUS
(B) THE TERM USAGE POSTING (''THE STATISTICS REPORT'')
(C) THE TERM USAGE LISTINGS
(D) THE THESAURUS OF ERIC DESCRIPTORS (''THE THESAURUS'')

D

SEARCHING FOR DOCUMENT CITATIONS USING A COMPUTER IS NOT DIFFICULT OR COMPLICATED. THERE ARE SOME GROUND RULES WHICH WILL MAKE YOUR COMPUTERIZED SEARCHING MORE FRUITFUL.

THE COMPUTER IS JUST A TOOL. TO USE IT WITH THE GREATEST AMOUNT OF PRODUCTIVITY AND THE LEAST AMOUNT OF DIFFICULTY, SOME SIMPLE CONCEPTS ABOUT COMPUTERS ARE NECESSARY.

VISUALIZE THE COMPUTER AS A ''BLACK BOX'' WITH A ''SLOT'' FOR TAKING IN THINGS (PROGRAMS, DATA, AND EVEN ENCODED SEARCH STRATEGIES) WHICH ARE NORMALLY CALLED ''INPUT''. SIMILARLY, OUR BOX HAS A PLACE FOR GETTING THINGS BACK (THE THINGS BEING ''OUTPUT'').

ALSO, CONSIDER THE FACT THAT WHAT GOES ON IN THE BLACK BOX IS OF LITTLE CONCERN TO US EXCEPT THAT WE SHOULD KNOW HOW IT WILL BEHAVE.

THE SYSTEM YOU WILL BE USING, CALLED QUERY, IS A SPECIAL PROGRAM WHICH WILL ALLOW YOU TO RETRIEVE INFORMATION FROM A VERY LARGE FILE OF CONDENSED DOCUMENT CITATIONS (KNOWN AS ''THE FILE'') -- PROVIDED THAT:

YOU PUT A CORRECT ENCODED SEARCH STRATEGY INTO THE ''INPUT SLOT'' AND YOU KNOW HOW THE BLACK BOX WILL BEHAVE TOWARD YOUR SEARCH.

WOULD YOU SAY THAT WHAT GOES ON IN THE ''BLACK BOX'' HAS BEEN PREDEFINED BY THE PROGRAM?

VIRTUALLY EVERYTHING NEEDED FOR YOUR
SPECIFIC SEARCH HAS BEEN PREDEFINED. THE NICE THING ABOUT THIS
IS THAT YOU DON'T HAVE TO WORRY ABOUT ANYTHING EXCEPT HOW TO
CREATE AN ENCODED SEARCH STRATEGY WHICH WILL BE ACCEPTED BY THE
BLACK BOX. (WE WILL GET INTO JUST HOW THIS IS DONE LATER!)

OK, HAVE YOU EVER DONE A COMPREHENSIVE SEARCH IN A LIBRARY?

YES
SEARCHING VIA A COMPUTER IS QUITE DIFFERENT FROM DOING
A LIBRARY SEARCH.

DO YOU KNOW WHAT THE DIFFERENCE IS?

NO
THE DIFFERENCE IS QUITE SIMPLE.
MOST PEOPLE DEFINE THEIR SEARCH NEEDS WHILE BROWSING
IN THE LIBRARY. SINCE BROWSING VIA COMPUTER IS NOT EASILY
DONE, YOU MUST DEFINE YOUR SEARCH PRIOR TO ASKING FOR
INFORMATION INSTEAD OF WHILE YOU ARE LOOKING FOR IT.

DO YOU SUPPOSE, BECAUSE OF THIS DIFFERENCE, THE PLANNING
WHICH YOU MUST DO IS MORE EXACT WHEN DOING A COMPUTER
SEARCH?

YES

TO CONTINUE, COPY (TYPE IN AND HIT THE RETURN BUTTON FOR EACH) THE FOLLOWING
SEQUENCE (ONE AT A TIME) EXACTLY.

```
)SAVE
)LOAD TWO
)COPY ONE DATA
START2



NOTICE THAT 'START2' DOES NOT HAVE THE ) TO THE LEFT OF IT.
BEGIN TYPING THE SEQUENCE.



      )SAVE
VALUE ERROR
      ) SAVE
        ^
      )SAVE
  30.07.26 06/21/72 ONE
      )SAVE)LOAD)COPYONEDATASTART2
INCORRECT COMMAND
      )SAVE
      )LOAD
           TWO

CHARACTER ERROR
      )
       ^

      )SAVE

      )LOAD TWO

      )COPY ONE DATA

      START2
RESEND

      )SAVE
  30.13.12 06/21/72 ONE

      )AV

INCORRECT COMMAND

      )SAVE

      )LOAD TWO
```

```
      )COPY ONE DATA

      START?
CHARACTER ERROR

      ^


      )LOAD TWO
SAVED  15.20.33 06/20/72
      )COPY ONE DATA
SAVED  20.13.12 06/21/72


      ST?
        ^
        ART?



DO YOU KNOW WHAT IS MEANT WHEN WE TALK ABOUT THE ''FILE''?



NO
THE ''FILE'', AS IT WILL BE CALLED HERE, REFERS TO A
PREDEFINED GROUP OF RECORDS CONSISTING OF HUNDREDS OF
DOCUMENT CITATIONS FROM WHICH YOU WILL WANT TO SELECT THE
ONES YOU NEED.

DO YOU KNOW WHAT A RECORD IN THE FILE LOOKS LIKE?



NO
A RECORD (WHICH REPRESENTS ONE DOCUMENT) CONSISTS OF A NUMBER OF
WORD GROUPINGS SUCH AS THE AUTHOR'S NAME, THE TITLE, AND THE
NUMBER OF PAGES CONTAINED IN THE ACTUAL COMPLETE DOCUMENT.

ARE YOU FAMILIAR WITH THE DESCRIPTOR FIELD WITHIN A RECORD?
```

NO
THE DESCRIPTOR FIELD CONTAINS, FOR SEARCH PURPOSES, THE MOST IMPORTANT
WORD GROUPS WITHIN A RECORD.
IT CONTAINS ALL OF THE DESCRIPTORS ASSIGNED TO THE
DOCUMENT AND IS USED TO LOCATE DOCUMENT CITATIONS. THUS,
IT IS NOT NECESSARY TO KNOW AN AUTHOR'S NAME OR THE TITLE
OF A DOCUMENT TO RETRIEVE ITS CITATION FROM THE FILE.

DO YOU KNOW WHAT DESCRIPTORS ARE?

NO
DESCRIPTORS ARE KEY WORDS ASSIGNED TO A DOCUMENT TO DESCRIBE ITS
CONTENT. DESCRIPTORS MAY BE SINGLE WORDS OR A GROUP OF WORDS.
FOR EXAMPLE, BUILDING IS A SINGLE WORD DESCRIPTOR AND
BUILDING DESIGN IS A MULTI-WORD DESCRIPTOR.

IMPORTANT
IT IS THE DESCRIPTOR WHICH IS USED TO FIND A CITATIONS IN THE FILE.

OK. WOULD YOU SAY THERE IS A RELATIONSHIP BETWEEN THE
DESCRIPTORS (KEY WORDS) YOU USE IN YOUR SEARCH AND WHAT
YOU GET BACK FROM THE COMPUTER?

YES
THE RETRIEVING OF DOCUMENTS FROM THE FILE DEPENDS ON THE
USE OF DESCRIPTORS.

MUST THE WHOLE DESCRIPTOR (SINGLE OR MULTI-WORD) BE USED?

YES
NOT REALLY. YOU MAY USE PARTS (WORDS) OF DESCRIPTORS, BUT
BE VERY CAREFUL WHEN YOU DO!

LOOK AT THESE DESCRIPTORS.

BUILDING CONVERSION
BUILDING DESIGN
MODULAR BUILDING DESIGN
BUILDING EQUIPMENT

IF YOU WANTED ALL DOCUMENTS RELATING TO BUILDINGS, WHICH DESCRIPTOR SHOULD YOU USE?
(TYPE IN THE DESCRIPTOR YOU THINK WOULD BE BEST)

BUILDING DESIGN
BY USING THE SINGLE WORD ''BUILDING''YOU ARE REALLY
SELECTING A GROUP OF DESCRIPTORS WHICH CONTAIN A COMMON WORD.
IF YOU USED OTHER WORDS IN ADDITION TO THE WORD ''BUILDING'',
JUST THE CITATIONS WITH THE MORE COMPLETE DESCRIPTOR WILL BE RETRIEVED.

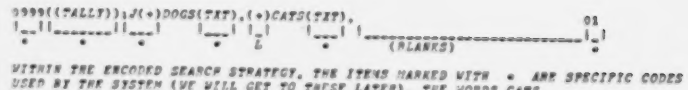
SO, FOR EXAMPLE, ''BUILDING DESIGN'' WILL RETRIEVE ALL
CITATION RECORDS WITH ''MODULAR BUILDING DESIGN'' AND
''BUILDING DESIGN'' FROM THE FILE. ''BUILDING
EQUIPMENT'' WILL RETRIEVE ONLY THOSE DOCUMENTS CONTAINING THAT
DESCRIPTOR. THE SAME IS TRUE OF THE DESCRIPTOR ''BUILDING CONVERSION''.

BUT,

JUST THE DESCRIPTOR ''BUILDING'' WILL RETRIEVED ALL
OF THE DOCUMENT RECORDS WITH ''BUILDING'' (THE COMMON WORD) IN THE
DESCRIPTOR FIELD.

IMPORTANT
EACH TIME YOU USE A DESCRIPTOR (SINGLE OR MULTI-WORD)
YOU MUST REMEMBER THAT THE COMPUTER WILL RETRIEVE
100 PERCENT OF THE DOCUMENT CITATIONS IN THE FILE WHICH
CONTAIN A CHOSEN DESCRIPTOR. IF, FOR EXAMPLE, YOU USE JUST
THE DESCRIPTOR 'ELECTR', OVER 1,000 DOCUMENTS WILL BE
RETRIEVED -- OBVIOUSLY, TOO MANY CITATIONS TO BE USEFUL.

THE INPUT TO THE COMPUTER (SYSTEM) CONSISTS OF IBM CARDS WHICH
TAKE THE FORM OF ENCODED SEARCH STRATEGIES. A MAJOR PORTION
OF THE TRAINING OFFERED HERE WILL CONSIST OF LEARNING HOW TO
CONSTRUCT A CORRECT SEARCH STRATEGY. BUT... FOR THE MOMENT
LETS LOOK AT WHAT AN ENCODED SEARCH STRATEGY LOOKS LIKE.



WITHIN THE ENCODED SEARCH STRATEGY, THE ITEMS MARKED WITH * ARE SPECIFIC CODES
USED BY THE SYSTEM (WE WILL GET TO THESE LATER). THE WORDS CATS
AND DOGS ARE THE DESCRIPTORS USED TO SEARCH THE FILE. THE ITEM
MARKED WITH L IS A LOGICAL OPERATOR (TO BE DISCUSSED LATER).

IT IS IMPORTANT TO REALIZE THAT IF ANY OF THE CODES MARKED WITH *
ARE OMITTED FROM THE ENCODED SEARCH STRATEGY, THE SYSTEM (THE COMPUTER)
WILL REJECT THE SEARCH.

DO YOU KNOW WHAT HAPPENS WHEN A SEARCH IS REJECTED?

NO
YOU GET NOTHING BACK! YOUR SEARCH WILL NOT RUN AND, TO GET THE
INFORMATION YOU WANT ANOTHER SEARCH WILL HAVE TO BE SUBMITTED.
THIS WASTES YOUR TIME AND COMPUTER FACILITIES.

OK. IT WILL BE BENEFICIAL FOR YOU TO REMEMBER THAT THE ENCODED
SEARCH STRATEGY WILL BE YOUR ONLY INPUT TO THE SYSTEM. THUS,
IT IS VITAL THAT YOU LEARN HOW TO CREATE SUCH.

SOME GROUND RULES:
BE EXACT.
DON'T BE AFRAID OF THE SYSTEM -- WE HAVEN'T LOST A CUSTOMER YET.
DON'T WORRY ABOUT THE SYSTEM. IT WILL DO ITS THING WITHOUT HELP.
DO CONCENTRATE UPON WHAT YOU GIVE THE SYSTEM TO WORK ON.
DO HAVE IN MIND WHAT YOU ARE LOOKING FOR. NO BROWSING.
HAVE FUN!

TYPE THE FOLLOWING SEQUENCE IN THE SAME MANNER AS PREVIOUSLY DONE.

)SAVE

)LOAD THREE

)COPY TWO DATA2

START3

)SAVE

)LOAD THREE

)COPY TWO DATA2

START3
CHARACTER ERROR

^

      v
      ART3
VALUE ERROR
      ART3
      ^
      ART3
VALUE ERROR
      ART3
      ^
)SAVE
  38.32.41 06/21/72 TWO
)LOAD THREE
SAVED 15.51.01 06/20/72
)COPY TWO DATA2
SAVED 38.32.41 06/21/72
      START3

TO DEFINE YOUR SEARCH, THREE
PUBLICATIONS ARE NEEDED:

THE THESAURUS OF ERIC DESCRIPTORS (ROTATED DESCRIPTOR
DISPLAY) -- HEREAFTER REFERRED TO AS THE 'ROTATED THESAURUS'.

THE ERIC DESCRIPTOR USAGE STATISTICAL REPORT -- HEREAFTER
CALLED THE 'DESCRIPTOR STATISTICS REPORT'.

THE THESAURUS OF ERIC DESCRIPTORS -- HEREAFTER CALLED
'THE THESAURUS'.

•••••••• THE THESAURUS ••••••••

DO YOU KNOW WHAT THE THESAURUS IS AND HOW TO USE IT?

NO
EACH DOCUMENT (OR CITATION) ENTERED INTO THE ERIC INFORMATION
HAS BEEN ''DESCRIBED'' BY A KNOWLEDGEABLE PERSON
USING 5 DESCRIPTORS. FOR EXAMPLE, A CITATION MIGHT BE 'DESCRIB.. WITH THE USE
OF 'PSYCHOLOGICAL TESTS', 'EDUCATIONAL TESTING',
'PRIMARY GRADES', AND OTHER RELEVANT DESCRIPTORS. THIS IS A
SHORTHAND WAY OF DESCRIBING THE CONTENTS OF A DOCUMENT.

THE ALLOWABLE DESCRIPTORS USED IN THIS PROCESS ARE
CONTAINED IN THE THESAURUS. THERE ARE OVER 4,700 DESCRIPTORS
FROM WHICH TO CHOOSE. THUS, THE THE THESAURUS IS JUST
A 'LIST' OF 'LEGAL' DESCRIPTORS WHICH MAY BE USED
TO IDENTIFY A DOCUMENT.

DO YOU KNOW WHY IT IS IMPORTANT FOR YOU TO BE ABLE TO USE
THE THESAURUS CORRECTLY?

NO
IF DOCUMENTS ARE DESCRIBED BY DESCRIPTORS FROM THE THESAURUS,
WOULD IT NOT BE POSSIBLE TO TRANSLATE YOUR INFORMATION NEEDS
INTO DESCRIPTORS ALSO?

YES
RIGHT! AND THAT IS EXACTLY WHAT YOU WILL NEED TO DO -- A
TASK WHICH WOULD BE IMPOSSIBLE UNLESS YOU HAD A LIST OF
'LEGAL' DESCRIPTORS TO USE.

•••IMPORTANT•••

THE KEY TO SUCCESSFUL RETRIEVAL IS KNOWING HOW TO USE THE
THESAURUS CORRECTLY!

DO YOU KNOW THE STEPS TO TRANSLATING YOUR INFORMATION NEED
(SEARCH) INTO THE CORRECT DESCRIPTORS?

NO
THE 1ST AND MOST IMPORTANT STEP IS FOR YOU TO STATE YOUR
INFORMATION NEED IN A BRIEF SENTENCE.

EXAMPLE: I NEED INFORMATION DEALING WITH THE USE OF
READING ACHIEVEMENT TESTING IN THE EARLY GRADES.

THE NEXT STEP IS TO IDENTIFY THE KEY WORDS IN YOUR
NEED STATEMENT.

EXAMPLE: I NEED INFORMATION DEALING WITH THE [USE] OF
[READING] [ACHIEVEMENT TESTING] IN THE [EARLY GRADES].

(SOMETIMES THE ''WHO, WHAT, WHEN, WHERE AND WHY ''
RULE IS HANDY FOR DOING THIS.)

THE NEXT TASK IS TO FIND THE CORRECT DESCRIPTORS IN THE
THESAURUS.

[USE]...(OF TESTS)
[ACHIEVEMENT TESTING]
[READING]
[EARLY GRADES]

MAY OR MAY NOT BE LEGAL DESCRIPTORS. TO DETERMINE
IF THEY ARE, YOU SHOULD LOOK THEM UP IN THE THESAURUS. (FIND IT NEXT
TO THE TERMINAL. HIT THE RETURN BUTTON WHEN YOU ARE READY.)

SO:
LOOKING ON PAGE 220 DO YOU FIND 'USE' OR 'USE OF TESTS'?

NO
'USE' OR 'USE OF TESTS' IS NOT THERE. IT MUST BE CONCLUDED
THAT THESE TERMS ARE NOT LEGAL.
USING THE THESAURUS AS A GUIDE, ENTER THE DESCRIPTOR YOU THINK WOULD BE A GOOD SUBSTITUTE.

USE STUDIES
A GOOD SUBSTITUTE MIGHT BE 'EDUCATIONAL TESTING' (WELL YOU CAN KNOW FOR SURE).

IS IT LEGAL?

OK. LET'S USE 'ACHIEVEMENT TESTS'.

AT THIS POINT, YOU MAY WANT TO DEFINE YOUR DESCRIPTOR IN
NARROWER OR BROADER TERMS.

NOTICE ON PG 2 THAT OTHER LEGAL DESCRIPTORS ARE GIVEN UNDER
'ACHIEVEMENT TESTS'. THE CODE 'NT' MEANS NARROWER TERM
AND 'BT' MEANS BROADER (MORE GENERAL) TERM.
ALSO, 'UF' MEANS ''USED FOR'' AND TELLS YOU NOT TO
USE A CERTAIN DESCRIPTOR(S). IN THIS CASE 'ACHIEVEMENT
PREDICTION' IS ILLEGAL AND THE UF IS TELLING YOU TO USE
'ACHIEVEMENT TESTS' INSTEAD.

RT APPEARS TO THE LEFT OF RELATED (RELATED TERM) DESCRIPTORS.

SOME ILLEGAL DESCRIPTORS (USUALLY COMMON TERMS) ARE LISTED
AND THE CODE 'USE' APPEARS BELOW TELLING YOU WHICH LEGAL
DESCRIPTOR TO USE IN ITS PLACE. (HIT RETURN AFTER LOOKING AT THESE.)

SOMETIMES 'SN' APPEARS UNDER DESCRIPTORS IN THE THESAURUS.
THIS MEANS ''SCOPE NOTE'' AND IS A SHORT DEFINITION OF THE TERM.

****IMPORTANT****
THE USE OF UF, USE, NT, BT, SN, AND RT ARE VERY
HANDY FOR TRANSLATING YOUR INFORMATION NEEDS INTO A GOOD
SEARCH STRATEGY.

WOULD YOU LIKE TO PRACTICE USING THE THESAURUS ONCE BEFORE CONTINUING?

YES
GOOD. HERE IS AN INFORMATIONAL NEED STATEMENT.

I WANT ALL INFORMATION AVAILABLE RELATING TO THE
USE OF PHOTOGRAPHS FOR TEACHING THE ENGLISH LANGUAGE TO
THIRD GRADE CHILDREN.

USING A PENCIL, PLACE BRACKETS AROUND THE KEYWORDS IN THE
ABOVE STATEMENT AND HIT THE RETURN BUTTON WHEN READY.

ALTHOUGH WE MAY NOT AGREE EXACTLY, YOU SHOULD HAVE:

I WANT ALL INFORMATION AVAILABLE RELATING TO THE

(USE OF PHOTOGRAPHS) FOR (TEACHING) THE (ENGLISH LANGUAGE) TO

(THIRD GRADE CHILDREN).

THE KEY WORDS TO BE TRANSLATED TO LEGAL DESCRIPTORS ARE:

(USE OF PHOTOGRAPHS)
(TEACHING)
(ENGLISH LANGUAGE)
 THIRD GRADE CHILDREN)

NOW, GO TO THE THESAURUS AND SEE IF YOU CAN FIND SOME
LEGAL DESCRIPTORS WITH WHICH YOU ARE SATISFIED. HIT THE
RETURN BUTTON WHEN YOU HAVE FOUND THEM.

AGAIN, WE MAY NOT AGREE EXACTLY BUT THE FOLLOWING DESCRIPTORS
ARE LEGAL AND WILL RETRIEVE HITS FOR THE SEARCH OF INTEREST.

PHOTOGRAPHS	PG. 149
TEACHING METHODS	A RELATED TERM ON PG. 208. TEACHING WAS TOO BROAD. SEE SCOPE NOTE UNDER TEACHING METHODS ON PG. 209.
ENGLISH	ENGLISH LANGUAGE COULD NOT BE FOUND BUT ENGLISH WAS FOUND ON PG. 67.
GRADE 3	THIRD GRADE CHILDREN COULD NOT BE FOUND BUT GRADE 3 (PG.65) OR
OR	PRIMARY GRADES (PG. 155) WILL GET THE JOB DONE.
PRIMARY GRADES	

WHEN READY TO CONTINUE, HIT RETURN.

YOU MAY NOT HAVE FOUND EXACTLY THE SAME DESCRIPTORS AS THOSE ABOVE.
IF YOU CAME CLOSE -- THAT IS GOOD ENOUGH! IT IS UNLIKELY THAT 2
PEOPLE WILL ARRIVE AT THE EXACT SAME SET OF SEARCH DESCRIPTORS.

********** THE USE OF THE ROTATED THESAURUS **********

DO YOU KNOW WHAT THE ROTATED THESAURUS IS AND HOW TO USE IT?

YES

•••••••••••••••• USING THE DESCRIPTOR STATISTICAL REPORT••••••••••••••••

DO YOU KNOW WHAT THE DESCRIPTOR STATISTICS REPORT IS AND HOW TO USE IT?

NO
YOU HAVE PROBABLY GUESSED THAT KNOWING THE NUMBER OF TIMES A
DESCRIPTOR HAS BEEN USED IS ESSENTIAL TO CREATING
A GOOD SEARCH.

THE DESCRIPTOR STATISTICS REPORT GIVES THIS INFORMATION.

DO YOU KNOW HOW TO READ IT?

NO
OK. IT IS QUITE SIMPLE.
LOOK AT THE TOP OF PG. 1 OF THE DESCRIPTOR STATISTICS REPORT
LOCATED NEAR THE TERMINAL. (HIT RETURN WHEN READY)

YOU SHOULD SEE:

DESCRIPTOR	TOTAL USAGE	ED	EP	ES	FR	OTHER	MAJOR USAGE	MINOR USAGE
ABBREVIATIONS	5	5					4	1

-FTC-

DO NOT BE CONCERNED WITH THE COLUMNS LABELED ED EP ES AND FR.
USE THE FOLLOWING ONLY:

DESCRIPTOR - A LEGAL DESCRIPTOR.
TOTAL USAGE - TOTAL NUMBER OF DOCUMENTS DESCRIBED BY THE DESCRIPTOR ON THE LEFT.
MAJOR USAGE - TOTAL NUMBER OF DOCUMENTS DESCRIBED AS MOST RELEVANT TO THE DESCRIPTOR TO THE LEFT.
MINOR USAGE - TOTAL NUMBER OF DOCUMENTS DESCRIBED AS SLIGHTLY RELEVANT TO THE DESCRIPTOR TO THE LEFT.

(ALSO, TOTAL USAGE = MINOR USAGE + MAJOR USAGE)

DO YOU KNOW WHEN TO USE THE TOTAL AND MAJOR USAGE COLUMNS?

NO
OK. WHEN YOU WANT A DESCRIPTOR TO BE USED AS A MAJOR DESCRIPTOR IN YOUR SEARCH USE ONLY THE MAJOR COLUMN.
IN THE SAME MANNER, US THE TOTAL USAGE COLUMN FOR THE DESCRIPTORS YOU WANT WHICH ARE TO INCLUDE BOTH MAJOR AND MINOR DESCRIPTORS. REMEMBER THAT MINOR DESCRIPTORS ALWAYS RETRIEVE BOTH MAJOR AND MINOR CITATIONS FROM THE FILES.

OK
WOULD YOU LIKE ADDITIONAL TRAINING ON ANY OF THE FOLLOWING?

(1) HOW TO USE THE THESAURUS.
(2) HOW TO USE THE ROTATED THESAURUS.
(3) HOW TO USE THE DESCRIPTOR STATISTICAL REPORT.

NO
ARE YOU READY TO CONTINUE?
YES
AT THIS TIME, YOU WILL BE ASKED TO ENTER THE DESCRIPTORS
WHICH SATISFY YOUR SEARCH NEEDS. YOU MAY USE UP TO 5
DESCRIPTORS AND MUST USE AT LEAST 2. MOST PEOPLE USE BETWEEN 2 AND 4!

****IMPORTANT****

BE SURE THE DESCRIPTORS YOU ENTER MEET YOUR INFORMATION NEED;
HAVE BEEN DERIVED USING THE PROPER STEPS; AND ARE LEGAL!

***** VERY IMPORTANT!!! *****

ALWAYS PLACE A STAR (*) DIRECTLY IN FRONT OF ALL DESCRIPTORS
WHICH SHOULD BE CONSIDERED AS MAJOR DESCRIPTORS BY THE SYSTEM!

FOR EXAMPLE: MAJOR DESCRIPTOR = *INDIANS
MINOR DESCRIPTOR = INDIANS

IF A * PRECEDES THE DESCRIPTOR IT WILL BE CONSIDERED A MAJOR DESCRIPTOR BY THE
COMPUTER. IF YOU WANT TO USE A DESCRIPTOR TO RETRIEVE DOCUMENT CITATIONS CONTAINING
EITHER A MAJOR OR MINOR DESCRIPTOR (I.E. TOTAL USAGE), DO NOT US THE *.

ENTER YOUR MOST IMPORTANT (MOST RELEVANT) DESCRIPTOR.

OPEN PLAN SCHOOLS
DESCRIPTOR NO. 1 IS OPEN PLAN SCHOOLS
ENTER YOUR SECOND MOST IMPORTANT DESCRIPTOR.

```
FLEXIBLE FACILITIES
DESCRIPTOR NO. 2 IS FLEXIBLE FACILITIES
ENTER YOUR NEXT MOST IMPORTANT DESCRIPTOR (HIT RETURN IF NONE.)


FLEXIBLE CLASSROOMS
DESCRIPTOR NO. 3 IS FLEXIBLE CLASSROOMS
ENTER YOUR NEXT MOST IMPORTANT DESCRIPTOR (HIT RETURN IF NONE).


DO YOU WANT TO CHANGE ANY OF THE DESCRIPTORS YOU ENTERED?


NO

YOUR ORDERED DESCRIPTORS ARE:

OPEN PLAN SCHOOLS     FLEXIBLE FACILITIES     FLEXIBLE CLASSROOMS


DO YOU WANT TO CHANGE ANY OF THE ABOVE DESCRIPTORS AT THIS TIME?


NO
COPY THE FOLLOWING SEQUENCE AS BEFORE.  IGNORE THE SYSTEM'S MESSAGES.


)SAVE

)LOAD FOUR

)COPY THREE DATA3

START4

BEGIN NOW.
      )SAVE
  39.04.37 06/21/72 THREE
      )LOAD FOUR
SAVED  14.29.48 06/21/72
      )COPY THREE DATA3
SAVED  39.04.37 06/21/72
     START4
```

********** SEARCH CONSTRUCTION **********

THERE ARE 2 PHASES TO ACCOMPLISHING THE SEARCH CONSTRUCTION TASK.
THESE ARE (1) ARRANGING YOUR LEGAL DESCRIPTORS INTO A LOGICAL ORDER,
AND (2) ENCODING THE SEARCH STRATEGY INTO A ''LANGUAGE''
ACCEPTABLE TO THE SYSTEM. NEITHER IS DIFFICULT!

PHASE 1 WILL BE COVERED AT THIS TIME AND THE LAST PHASE
WILL BE COMPLETED SUBSEQUENTLY.

OK, NOW LET'S INSERT SOME LOGIC OPERATORS INTO THE LINE OF YOUR
DESCRIPTORS.
THE LOGIC OPERATOR (+) MEANS 'AND' AND (/) MEANS 'OR'
SUCH THAT:

*** THIS IS IMPORTANT! ***
CATS (+) DOGS MEANS ''RETRIEVE ALL CITATIONS WHICH HAVE
BOTH 'CATS' AND 'DOGS' ENTERED IN THE DESCRIPTOR FIELD
OF THE RECORD.

CATS (/) DOGS MEANS ''RETRIEVE ALL CITATIONS WHICH HAVE
EITHER 'CATS' OR 'DOGS' IN THE DESCRIPTOR FIELD. IF
BOTH DESCRIPTORS ARE CONTAINED ON A CITATION, IT WILL BE ALSO
RETRIEVED.

USING YOUR DESCRIPTORS, ONE POSSIBLE USE OF THE LOGIC OPERATOR
'(+)' IS:

OPEN PLAN SCHOOLS (+) FLEXIBLE FACILITIES (+) FLEXIBLE CLASSROOMS

NOTICE THAT THE LOGIC OPERATOR(S) IN THE ABOVE ARE '(+)'.

THUS, ALL DESCRIPTORS MUST BE CONTAINED IN A CITATION
BEFORE IT WILL BE RETRIEVED. THIS MEANS FEWER CITATIONS
MAY POSSIBLY BE RETRIEVED SINCE FEWER WILL CONTAIN ALL
THE DESCRIPTORS.

ON THE OTHER EXTREME:

OPEN PLAN SCHOOLS (/) FLEXIBLE FACILITIES (/) FLEXIBLE CLASSROOMS

HAS '(/)' LOGIC OPERATOR(S). THIS MEANS THAT CITATIONS
CONTAINING ANY OF THE DESCRIPTORS WILL BE RETRIEVED. THUS,
A LARGE NUMBER OF CITATIONS WILL BE RETRIEVED FROM THE
SYSTEM.

***** IMPORTANT *****

THERE WILL BE VERY FEW OCCASIONS WHEN ALL '(/)' LOGIC
OPERATORS ARE NECESSARY FOR AN INDIVIDUAL SEARCH.

ANOTHER POSSIBLE LOGICAL OPERATOR ARRANGEMENT IS:

OPEN PLAN SCHOOLS (+) FLEXIBLE FACILITIES (/) FLEXIBLE CLASSROOMS

USING LETTERS TO REPRESENT YOUR DESCRIPTORS, THE FOLLOWING
LOGIC OPERATOR COMBINATIONS ARE POSSIBLE.

[1] A (+) B (+) C

[2] A (+) B (/) C

[3] A (/) B (+) C

[4] A (/) B (/) C

SELECT FROM THE POSSIBILITIES ABOVE, THE LOGIC ORDER
YOU FEEL WOULD BE APPROPRIATE FOR YOUR SEARCH. ENTER
THE LINE NUMBER (NO BRACKETS NEEDED) FOR A MORE SPECIFIC
EXPLANATION OF WHAT CITATIONS WILL BE RETRIEVED BY YOUR CHOICE.

2
ANY CITATION WITH (A AND B) OR (A AND C) WILL BE RETRIEVED BY THE SYSTEM.

NOTICE THAT THE RESULTS OF A LOGIC PATTERN MAY BE DETERMINED BY READING THE LINE
FROM LEFT TO RIGHT. THIS WILL BE TRUE IN ALL CASES.
IF YOU WOULD LIKE TO EXAMINE ANOTHER, ENTER THE LINE NUMBER FROM ABOVE.
IF NOT, HIT RETURN.

LOGIC FOR DESCRIPTOR STATISTICAL REPORT AS YOUR COUNCE, ENTER
THE NUMBER OF TIMES EACH DESCRIPTOR WAS BEEN USED.

***** IMPORTANT *****

IF YOU USE A MODIFIED DESCRIPTOR, ALWAYS INCLUDE THE SUM OF THE
DESCRIPTORS REPRESENTED AS THE USAGE TOTAL.

USE THE MAJOR USAGE COLUMN IF YOU HAVE DESIGNATED A DESCRIPTOR AS
"MAJOR". OTHERWISE USE THE TOTAL USAGE COLUMN.

ENTER THE NUMBER OF TIMES 'OPEN PLAN SCHOOLS' HAS BEEN USED.
□:
20
ENTER THE NUMBER OF TIMES 'FLEXIBLE FACILITIES' HAS BEEN USED.
□:
54
ENTER THE NUMBER OF TIMES 'FLEXIBLE CLASSROOMS' HAS BEEN USED.
□:
16

USING THE FIGURES YOU ENTERED, THE FOLLOWING IS A LISTING OF
THE POSSIBLE LOGIC COMBINATIONS AND THE ESTIMATED NUMBER
OF CITATIONS WHICH MIGHT BE RETRIEVED.

OPEN PLAN SCHOOLS (*) FLEXIBLE FACILITIES (*) FLEXIBLE CLASSROOMS WILL RETRIEVE ABOUT 1 CITATION(S).

OPEN PLAN SCHOOLS (*) FLEXIBLE FACILITIES (/) FLEXIBLE CLASSROOMS WILL RETRIEVE ABOUT 1 CITATION(S).

OPEN PLAN SCHOOLS (/) FLEXIBLE FACILITIES (*) FLEXIBLE CLASSROOMS WILL RETRIEVE ABOUT 1 CITATION(S).

OPEN PLAN SCHOOLS (/) FLEXIBLE FACILITIES (/) FLEXIBLE CLASSROOMS WILL RETRIEVE NO FEWER THAN 20 AND UP TO 100 CITATIO
S).

THE ABOVE ARE ESTIMATES BASED ON CLASSICAL PROBABILITY THEORY
AND ARE NOT THE ACTUAL NUMBER OF CITATIONS YOU MIGHT RETRIEVE.
IN SHORT -- THEY ARE ESTIMATES ONLY.

DO YOU WANT TO CHANGE YOUR PREVIOUSLY SELECTED DESCRIPTORS IN ORDER TO IMPROVE THE NUMBER OF
CITATIONS YOU MIGHT RECEIVE FROM YOUR SEARCH?

NO

DO YOU WANT TO (RE)EXAMINE THE LOGIC ARRANGEMENT OR ESTIMATED HITS FOR YOUR DESCRIPTORS?
NO
TYPE THE FOLLOWING SEQUENCE AS BEFORE. IGNORE THE SYSTEM'S MESSAGES.

)SAVE

)LOAD FIVE

)COPY FOUR DATA↵

```
STARTS

*BEGIN NOW.
      )SAVE
  39.15.06 06/21/72 FOUR
      )LOAD FIVE
SAVED  39.45.59 06/21/72
      )COPY FOUR DATA*
SAVED  39.15.06 06/21/72
      STARTS


YOU ARE NOW ENTERING THE FINAL PHASE OF THIS TRAINING PROGRAM.

TO BE MORE SPECIFIC, THIS IS THE ''NITTY GRITTY'' OF THE SESSION.


********** ENCODING THE SEARCH STRATEGY **********


AN EXAMPLE OF A PROPERLY ENCODED SEARCH IS SHOWN BELOW.

      COLUMN 10          COLUMN 20          COLUMN 30          COLUMN 70
          *                  *                  *                  *
          1         2        3        *         5         6        7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890

0001((TALLY));J(*)*CATS(TXT),(/)*DOGS(TXT).                                    01

THE BLOCK OF NUMBERS ABOVE THE LINE BEGINNING '0001((TALLY));J( -ETC.- ' REPRESENTS
THE COLUMNS OF A STANDARD IBM CARD. NOTICE THAT THE SEARCH IS PRINTED AS IF IT WERE
BEING PUNCHED ON AN IBM CARD.  THIS IS NECESSARY SO THAT YOUR SEARCH CAN BE READ BY
THE COMPUTER. THE BLOCK OF NUMBERS CORRESPONDS TO CARD COLUMNS. NOTICE THAT THEY
START AT COLUMN 1 AND CONTINUE TO COLUMN 80.

LET'S TAKE A CLOSER LOOK AT THE EXAMPLE.
          1         2        3        *         5         6        7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890

0001((TALLY));J(*)*DOGS(TXT),(*)*CATS(TXT).
|__||----------||__||___||_||__||___||----------------------------------|_||
 A      B        C   D    E  F   G                  H                   I  01

WHERE:
A    IS THE SEARCH NUMBER WHICH YOU FURNISH. IT MAY BE ANY 4 DIGITS AND MUST START IN COLUMN 1.
B    IS THE 'TALLY GROUP' WHICH MUST ALWAYS BE ON THE 1ST CARD OF A SEARCH AND START IN COLUMN 5.
C    IS YOUR FIRST DESCRIPTOR. IT MUST START IN COLUMN  10 OF THE 1ST CARD BUT WILL VARY IN LENGTH
     DEPENDING ON THE LENGTH OF THE DESCRIPTOR YOU CHOOSE.
D    IS A 'TEXT GROUP' WHICH MUST FOLLOW EVERY DESCRIPTOR USED IN THE SEARCH.
E    IS A LOGIC OPERATOR. LOGIC OPERATORS MUST APPEAR ONLY BETWEEN DESCRIPTORS.
F    IS ANOTHER DESCRIPTOR (IN THIS CASE THE LAST).
G    IS ANOTHER ' TXT GROUP' (ALSO THE LAST IN THIS EXAMPLE).
H    ARE BLANK COLUMNS, I.E. SPACES NOT NEEDED FOR THIS PARTICULAR SEARCH. *** NOTICE*** THE ONLY
     BLANKS ALLOWED IN YOUR SEARCH ARE THOSE WHICH SHOULD APPEAR BETWEEN WORDS OF DESCRIPTORS AND
     THOSE WHICH COME AFTER THE LAST DESCRIPTOR AND 'TXT' GROUP.  NO OTHER BLANKS ARE ALLOWED!
```

NOTICE THAT THE FOLLOWING TWO CARD SEARCH HAS NO MULTIPLE BLANKS ON THE FIRST
(TO BE CONTINUED) CARD. DO NOT LEAVE MULTIPLE BLANKS IN YOUR SEARCH IF YOU
MUST CONTINUE TO A SECOND (OR THIRD) CARD!

```
0001((TALLY));J(+)*PHOTOGRAPHY(TXT),(+)*AUDIOVISUAL INSTRUCTION(TXT),(+)*AUTOI01
0001#STRUCTIONAL METHODS(TXT),(+)*ADULT STUDENTS(TXT).                      02
```

T IS THE CARD NUMBER. *** NOTICE *** OFTEN, SEARCHES ARE TOO LONG
TO BE PLACED ON A SINGLE CARD. WHEN THIS SITUATION OCCURS, A 2ND
OR 3RD CARD MAY BE USED TO COMPLETE THE SEARCH. ENTER 01, 02 OR 03 IN COLUMNS 79 - 80
DEPENDING WHICH 'CARD' YOU ARE USING.

TO ENTER A SEARCH, YOU NEED ONLY TO TYPE THE SEQUENCE GIVEN ABOVE ON THE
'CARD' WHEN REQUESTED TO DO SO.

***** IMPORTANT *****

IT WILL BE YOUR RESPONSIBILITY TO ENTER ALL PARTS OF THE SEARCH
STRATEGY IN THE PROPER ORDER WHILE OBSERVING THE PROPER CARD
COLUMNS.

YOUR DESCRIPTORS ARE (BE SURE THEY ARE LEGAL) :

OPEN PLAN SCHOOLS FLEXIBLE FACILITIES FLEXIBLE CLASSROOMS

AT THIS TIME, ENTER THE FIRST 'CARD' OF YOUR SEARCH.
DO NOT PUSH THE RETURN BUTTON UNTIL YOU ARE SATISFIED WITH YOUR ENTRY.
ALSO MAKE SURE YOUR 'CARD' IS 80 CHARACTERS LONG. (DON'T FORGET TO INCLUDE THE CARD NUMBER!)
*** NOTICE *** CORRECTIONS ARE MADE IN THE USUAL MANNER.

```
        1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
```

```
1234((TALLY));J(+)OPEN PLAN SCHOOLS(TXT),(+)FLEXIBLE FACILITIES(TXT),(/)FLEX+ 01
                                                                          v
                                                                          IB01
```

WAS YOUR SEARCH TOO LONG FOR ONE 'CARD'?
YES
OK. ENTER THE 2ND 'CARD' OF YOUR SEARCH. *** NOTICE *** ONLY
THE SEARCH NUMBER, YOUR CONTINUATION, AND THE CARD NUMBER (COLUMNS 79-80 SHOULD BE '02')
ARE REQUIRED. IN SHORT, THE 'TALLY GROUP' DOES NOT GO ON THIS 2ND 'CARD'.

```
        1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
1234+LE(TXT),
```

```
     v
     CLASSROOMS(TXT).
                                                                          02
```

```
IS YOUR SEARCH TOO LONG FOR 2 'CARDS'?
NO
DO YOU WANT YOUR SEARCH CARDS LISTED AT THIS TIME?
YES

        1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
1234((TALLY));J(+)OPEN PLAN SCHOOLS(TXT),(+)FLEXIBLE FACILITIES(TXT),(/)FLEXIB01
1234LE CLASSROOMS(TXT).
***** NO CARD 3 ENTERED *****                                                  02

DO YOU WANT TO CHANGE ANY OF YOUR CARDS?
NO


YOUR SEARCH IS NOW BEING SCANNED FOR ERRORS.

NO THIRD SEARCH CARD FOUND. IS THIS OK?
YES
IT WILL TAKE ABOUT ONE MINUTE TO COMPLETE THE SCANNING PROCESS.

******************************
* YOUR SEARCH MAY  NEED TO   *
* BE CORRECTED. A LIST OF    *
* POSSIBLE ERRORS FOLLOWS.   *
******************************

[[ RESULTS OF DIAGNOSTICS ]]
**WARNING ONLY** - NO MAJOR DESCRIPTORS FOUND.
[[ END OF DIAGNOSTIC MESSAGES ]]


YOUR SEARCH LOOKS LIKE THIS. (IGNORE THE CARDS YOU DID NOT USE)
        1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
1234((TALLY));J(+)OPEN PLAN SCHOOLS(TXT),(+)FLEXIBLE FACILITIES(TXT),(/)FLEXIB01  <CARD NO. 1
1234LE CLASSROOMS(TXT).                                                         02  <CARD NO. 2
1234                                                                            03  <CARD NO. 3

DO YOU NEED TO CHANGE ANY OF THEM?
YES
THE FIRST, THE SECOND, THE THIRD, OR ALL OF THEM?
FIRST
        1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
1234((TALLY));J(+)OPEN PLAN SCHOOLS(TXT),(+)FLEXIBLE FACILITIES(TXT),(/)FLEXI01

IT WILL TAKE ABOUT ONE MINUTE TO COMPLETE THE SCANNING PROCESS.

*** DIAGNOSTICS COMPLETED.  NO ERRORS FOUND. ***


YOUR SEARCH LOOKS LIKE THIS. (IGNORE THE CARDS YOU DID NOT USE)
        1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
1234((TALLY));J(+)OPEN PLAN SCHOOLS(TXT),(+)FLEXIBLE FACILITIES(TXT),(/)FLEXI01  <CARD NO. 1
1234LE CLASSROOMS(TXT).                                                         02  <CARD NO. 2
1234                                                                            03  <CARD NO. 3

DO YOU NEED TO CHANGE ANY OF THEM?
YES
THE FIRST, THE SECOND, THE THIRD, OR ALL OF THEM?
```

```
SECOND
REENTER IT AT THIS TIME.
          1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
1234BLE CLASSROOMS(TXT).                                                      02
IT WILL TAKE ABOUT ONE MINUTE TO COMPLETE THE SCANNING PROCESS.

*** DIAGNOSTICS COMPLETED.  NO ERRORS FOUND. ***


YOUR SEARCH LOOKS LIKE THIS. (IGNORE THE CARDS YOU DID NOT USE)
          1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
1234((TALLY))↓J(+)←OPEN PLAN SCHOOLS(TXT),(+)FLEXIBLE FACILITIES(TXT),(/)FLEXI01  ←CARD NO. 1
1234←BLE CLASSROOMS(TXT).                                                     02  ←CARD NO. 2
1234                                                                          03  ←CARD NO. 3


DO YOU NEED TO CHANGE ANY OF THEM?
NO
ARE YOU READY TO SUBMIT YOUR SEARCH? (DO NOT SAY YES IF ERRORS ARE NOTED ABOVE.)
YES


ENTER YOUR NAME.

ENTER A PHONE NUMBER WHERE YOU CAN BE REACHED WHEN YOUR SEARCH IS READY.
646-2115
ENTER  A SHORT STATEMENT OF YOUR INTENDED USE OF THE SEARCH RESULTS.
I INTEND TO USE THE RESULTS FOR TEACHING AND PURPOSES.

                                                  ▼
                                                  WRITING PURPOSES.
IF YOUR SEARCH IS TO BE USED FOR A CLASS PROJECT, ENTER THE COURSE NUMBER
AND/OR COURSE NAME.
ED. 4663 INNOVATIONS IN EDUCATION


YOUR SEARCH HAS BEEN SUBMITTED!


ENTER THE FOLLOWING SEQUENCE AS BEFORE. IGNORE THE SYSTEM'S MESSAGES. START AFTER THE 'SENT' MESSAGE.

)SAVE

)LOAD SIX

)COPY FIVE DATA5

START6


ADL32979 SENT
      )SAVE
  40.05.43 06/21/72 FIVE
      )LOAD SIX
SAVED  15.56.04 06/21/72
      )COPY FIVE DATA5
SAVED  40.05.43 06/21/72
```

STARTS

WHEN YOU FIRST STARTED THIS TRAINING SESSION, YOU WERE ASKED TO ANSWER 16 QUESTIONS WHICH WERE USED TO SET UP YOUR TRAINING.

NOW THAT YOU HAVE COMPLETED THE SESSION, PLEASE ANSWER THE SAME QUESTIONS AGAIN. THIS TIME, AFTER ANSWERING EACH QUESTION, YOU WILL BE INFORMED AS TO THE CORRECTNESS OF BOTH ANSWERS, YOUR FIRST AND YOUR LAST.

ALSO, IMMEDIATELY AFTER ANSWERING THE LAST QUESTION, YOUR ''SCORES'' FOR BOTH QUESTION SECTIONS, AND THE DIFFERENCE BETWEEN THEM, WILL BE COMPUTED FOR YOU.

IT IS EMPHASIZED THAT THE ''SCORES'' ARE REALLY NOT SCORES AS SUCH, BUT A SIMPLE METHOD OF LETTING YOU COMPARE YOUR EFFORTS.

◆◆◆ READ THESE INSTRUCTIONS CAREFULLY. ◆◆◆

ΔΔΔΔΔΔΔΔΔΔΔΔΔΔ
Δ THE FOLLOWING QUESTIONS ARE OF 2 TYPES, TRUE OR FALSE, AND MULTIPLE CHOICE. ANSWER THE TRUE OR FALSE
Δ QUESTIONS WITH THE WORDS TRUE, OR FALSE. (DO NOT USE 'YES' OR 'NO'.) ANSWER MULTIPLE CHOICE
Δ QUESTIONS BY SELECTING THE APPROPRIATE LETTER OR NUMBER.
▼▼▼▼▼▼▼▼▼▼▼▼▼▼

1.
CITATIONS OF DOCUMENTS ANNOUNCED IN BOTH RESEARCH IN EDUCATION (RIE) AND CURRENT INDEX TO JOURNALS IN EDUCATION (CIJE) CAN BE RETRIEVED BY THE COMPUTER?

TRUE
CORRECT!
CORRECT THE 1ST TIME.

2.
SEARCHING FOR INFORMATION BY COMPUTER IS MORE EXACT (MUST BE BETTER PLANNED) THAN SEARCHING IN A LIBRARY?

TRUE
CORRECT!
CORRECT THE 1ST TIME.

3.
TO OBTAIN RELEVANT INFORMATION FROM THE COMPUTER, YOU MUST KNOW THE AUTHOR'S NAME OR THE TITLE OF THE DOCUMENT DESIRED?

FALSE
CORRECT!
CORRECT THE 1ST TIME.

4.
A COMPUTER RETRIEVED DOCUMENT THAT YOU FEEL IS VERY RELEVANT TO YOUR INFORMATION NEED IS CALLED A ''HIT''. TO OBTAIN A SMALL NUMBER OF HITS FROM A SEARCH STRATEGY, IT IS BETTER TO DEFINE YOUR SEARCH IN BROAD TERMS (THAT IS, USE DESCRIPTORS WHICH FIT A LARGER CATEGORY OF DOCUMENTS)?

TRUE
INCORRECT.
INCORRECT THE 1ST TIME.

5.
WHICH OF THE FOLLOWING DESCRIPTORS WOULD BE MOST APPROPRIATE FOR SEARCHING THE FILE FOR DOCUMENTS RELATING TO ''THE TEACHING OF DISADVANTAGED FIRST GRADE MEXICAN AMERICANS''?
[1] BILINGUAL EDUCATION (AND) GRADE 1
[2] BILINGUAL (AND) GRADE 1
[3] DISADVANTAGED (AND) MEXICAN AMERICANS (AND) TEACHING (AND) GRADE 1
[4] BILINGUAL EDUCATION (AND) GRADE 1 (AND) MEXICAN AMERICANS
[5] DISADVANTAGED
[6] MEXICAN AMERICANS (AND) TEACHING

3
CORRECT!
CORRECT THE 1ST TIME.

6.
WHICH OF THE FOLLOWING IS NOT A NECESSARY TOOL FOR FORMULATING A COMPUTER SEARCH?
[1] THE ROTATED THESAURUS
[2] THE THESAURUS OF ERIC DESCRIPTORS (''THE THESAURUS'')
[3] RESEARCH IN EDUCATION (RIE)
[4] THE TERM USAGE POSTINGS (''THE STATISTICS REPORT'')

3
CORRECT!
INCORRECT THE 1ST TIME.

7.
(+) AND (/) ARE THE PROPER SYMBOLS FOR THE LOGICAL
OPERATORS 'AND' AND 'OR'?

TRUE
CORRECT!
INCORRECT THE 1ST TIME.

8.
IN THE THESAURUS OF ERIC DESCRIPTORS, THE CODES ''RT'', ''SN'', ''BT'',
''UF'', ''NT'', AND ''USE'' MAY BE FOUND UNDER THE DESCRIPTORS. WHICH
OF THE FOLLOWING IS AN INCORRECT DESCRIPTION FOR THE CODE GIVEN?
(1) RT = RELATED TERM
(2) SN = SINGLE NOUN
(3) BT = BROADER TERM
(4) UF = USE FOR
(5) NT = NARROWER TERM
(6) USE = USE THIS TERM

2
CORRECT!
INCORRECT THE 1ST TIME.

9.
WHAT PART OF THE DOCUMENT RECORD (OR THE COMPUTER FILE) IS NORMALLY SEARCHED?
(1) THE AUTHOR'S NAME SECTION
(2) THE TITLE SECTION
(3) THE DESCRIPTOR SECTION
(4) THE ACCESSION NUMBER (''ED NUMBER'' OR ''EJ NUMBER'')

3
CORRECT!
CORRECT THE 1ST TIME.

10.
IF YOU SUBMIT A MECHANICALLY INCORRECT SEARCH,
(A) THE COMPUTER WILL BLOW UP
(B) YOUR SEARCH WILL RUN BUT WILL BE INVALID
(C) YOUR SEARCH WILL NOT RUN AND NO LISTINGS (RETRIEVED DOCUMENTS) WILL BE PRODUCED
(D) YOUR SEARCH WILL RUN BUT WILL TAKE TOO MUCH COMPUTER TIME
(E) THERE IS NO WAY OF KNOWING WHAT WILL HAPPEN

B
INCORRECT.
INCORRECT THE 1ST TIME.

11.
DESCRIPTORS MUST BE ENTERED ON YOUR SUBMITTED SEARCH
EXACTLY AS LISTED IN THE THESAURUS?

TRUE
INCORRECT
INCORRECT THE 1ST TIME.

12.
IN COMPUTER SEARCHING, A DESCRIPTOR NOT PRECEDED BY A STAR (*) WILL
RETRIEVE ALL DOCUMENT CITATIONS CONTAINING THAT PARTICULAR DESCRIPTOR --
REGARDLESS OF WHETHER MAJOR OR MINOR STATUS WAS ASSIGNED TO THAT DESCRIPTOR?

TRUE
CORRECT!
CORRECT THE 1ST TIME.

13.
THE COMPUTER WILL RETRIEVE 100 PERCENT OF THE DOCUMENTS
CONTAINING A PARTICULAR DESCRIPTOR?

FALSE
INCORRECT.
INCORRECT THE 1ST TIME.

THE FOLLOWING ARE SOME SEARCH STRATEGIES TO BE USED IN SUBSEQUENT QUESTIONS (ASSUME CORRECT CARD COLUMNS).

```
[A] 9999((TALLY))↓J(*)CATS(TXT),(*)DOGS(TXT),(*)MICE(TXT).                  01
[B] 9999((TALLY))↓J(*)CATS(TXT),(/)DOGS(TXT),(/)MICE.                       01
[C] 9999((TALLY))↓J(*)CATS(TXT),(*)DOGS.                                    01
[D] 9999((TALLY))↓J(*)CATS(TXT)                                             01
```

14.
OF THE FOUR ENCODED SEARCH STRATEGIES GIVEN ABOVE,
WHICH IS CORRECTLY ENCODED? (A, B, C, OR D)

A
CORRECT!
INCORRECT THE 1ST TIME.

15.
WHICH, IF IT WERE PROPERLY ENCODED, WOULD RETRIEVE THE LARGEST NUMBER OF HITS?
(A, B, C, OR D)

B
CORRECT!
INCORRECT THE 1ST TIME.

16.
WHICH, IF IT WERE PROPERLY ENCODED, WOULD RETRIEVE THE FEWEST NUMBER OF HITS?
(A, B, C, OR D)

D
INCORRECT.
INCORRECT THE 1ST TIME.

17.
TO RETRIEVE THE DESIRED RESULTS FROM A COMPUTER SEARCH,
WHICH IS (ARE) THE MOST IMPORTANT CONSIDERATION(S)?
(A) KEEPING THE CHARACTERS IN PROPER CARD COLUMNS.
(B) SELECTING DESCRIPTORS.
(C) SELECTING THE LOGIC ORDER.
(D) ALL OF THE ABOVE.
(E) NONE OF THE ABOVE.

D
CORRECT!
CORRECT THE 1ST TIME.

18.
IF IN DOUBT ABOUT HOW A DESCRIPTOR HAS BEEN DEFINED, IT WOULD BE TO YOUR ADVANTAGE TO
REFER TO:
(A) THE ROTATED THESAURUS
(B) THE TERM USAGE POSTING (" THE STATISTICAL REPORT")
(C) THE TERM USAGE LISTINGS
(D) THE THESAURUS OF ERIC DESCRIPTORS ("THE THESAURUS")

A
INCORRECT.
CORRECT THE 1ST TIME.

```
YOUR SCORE FOR THE FIRST TIME YOU ANSWERED THE QUESTIONS WAS:

48
YOUR SCORE THIS TIME IS:

57
WHICH IS A DIFFERENCE OF 22 POINTS.




ENTER THE FOLLOWING SEQUENCE AS DONE BEFORE.

)SAVE

)LOAD SEVEN

)COPY SIX DATA6

START?


BEGIN.
      )SAVE
  40.18.26 06/21/72 SIX
      )LOAD SEVEN
SAVED  15.56.34 06/21/72
      )COPY SIX DATA6
SAVED  40.18.26 06/21/72
      START?
THANK YOU                FOR TAKING YOUR VALUABLE TIME TO LEARN SOMETHING NEW.
COME BACK TO VISIT OFTEN!

YOU HAVE JUST BEEN SIGNED OFF.
PLEASE LEAVE YOUR TERMINAL OK.

4:19 PM WEDNESDAY  JUNE 21, 1972

  40.19.12 06/21/72 SEVEN
```

SUBJECT NO. 26 NAME:

DATE: 71272

BLOCK 1 (SYSTEM FAMILIARIZATION) REAL TIMES:

TERMINAL TRAINING	2	8
KEYBOARD TRAINING	4	32
RESPONSE TRAINING	1	31
SYSTEMS TRAINING	2	43
COMPUTER INPUT TRAINING	1	47
COMPUTER OUTPUT TRAINING	0	0
TOTAL	10	161

BLOCK 1 (SYSTEM FAMILIARIZATION) KEYBOARD WAIT TIMES:

TERMINAL TRAINING	0	29
KEYBOARD TRAINING	0	36
RESPONSE TRAINING	0	26
SYSTEMS TRAINING	0	21
COMPUTER INPUT TRAINING	0	1
COMPUTER OUTPUT TRAINING	0	0
TOTAL	0	113

BLOCK 2 (QUERY FAMILIARIZATION) REAL TIMES:

FILE TRAINING	4	57
ROTATED THESAURUS	56	4
DESCRIPTOR STATISTICS	2	21
THESAURUS	1	6
TOTAL	63	88

BLOCK 2 (QUERY FAMILIARIZATION) KEYBOARD WAIT TIMES:

FILE TRAINING	0	39
ROTATED THESAURUS	0	18
DESCRIPTOR STATISTICS	0	11
THESAURUS	0	5
TOTAL	0	73

BLOCK 3 (SEARCH STRATEGIES) REAL TIMES:

SELECTING DESCRIPTORS	21	0
LOGIC TRAINING	7	18
ESTIMATED HITS	10	1
TOTAL	38	19

BLOCK 3 (SEARCH STRATEGIES) KEYBOARD WAIT TIMES:

SELECTING DESCRIPTORS	18	17
LOGIC TRAINING	3	5
ESTIMATED HITS	5	8
TOTAL	26	30

BLOCK 4 (PRACTICE AND SUBMISSION) REAL TIMES:

ENCODING THE SEARCH	13	35
DIAGNOSTICS AND REWORK	2	28
TOTAL	15	63

BLOCK 4 (PRACTICE AND SUBMISSION) KEYBOARD WAIT TIMES:

ENCODING THE SEARCH	6	20
DIAGNOSTICS AND REWORK	0	40
TOTAL	6	60

TESTING RECORD

REAL TIMES:

PRE-TEST	9	45
POST-TEST	10	31
TOTAL	19	76

KEYBOARD WAIT TIMES:

PRE-TEST	2	10
POST-TEST	2	24
TOTAL	4	34

SCORES:

PRE-TEST: 39
POST-TEST: 84
DIFFERENCE: 45

RESPONSES:

28, PG. 3

1	1	1
2	1	1
3	1	1
4	1	1
5	1	1
6	0	1
7	0	1
8	0	1
9	0	1
10	0	1
11	0	0
12	0	1
13	0	1
14	0	0
15	0	0
16	1	1
17	1	1
18	0	1

MAJOR TOTALS:

TOTAL TIME ON TERMINAL	34	39	
TOTAL KEYBOARD WAIT TIME	17	33	
TOTAL CPU TIME	0	20	32

SEARCH:

```
????((TALLY));J(+)BILINGUAL(TXT),(+)MEDIATIONAL THEORY(TXT),(/)CULTURE FREE TE01
????STS(TXT),(/)LEARNING PROCESSES(TXT),(/)CONCEPT FORMATION(TXT),          02
????                                                                        03
```

NOTES:

10:45 AM WEDNESDAY JULY 12, 1972

Appendix C

Instruments and Forms

COMPUTER SEARCH INPUT INFORMATION

NAME:_____.____. CLASSIFICATION:
(LAST) (FIRST) ☐ JR ☐ SR
GRAD: ☐ M
MAJOR:_____. ☐ SP ☐ DR

COLLEGE:_____. UCC CHARGE CODE:_____.

DATE:____|____|72 PHONE:_____. SEX: ☐ M ☐ F

USING 1 TO 2 SENTENCES, GIVE YOUR SPECIFIC INFORMATIONAL NEED:

_____

_____

_____

_____.

INDICATE YOUR SEARCH PREFERENCE [CHECK ONE]:

☐ WANT ALL DOCUMENT CITATIONS TO BE RETRIEVED BY THE SYSTEM REGARDLESS OF HOW RELEVANT (I.E. HIGHLY RELEVANT AS WELL AS SLIGHTLY RELEVANT) SUCH CITATIONS MIGHT BE TO THE INFORMATIONAL NEED STATED PREVIOUSLY.

☐ WANT ALL OF THE DOCUMENT CITATIONS WHICH ARE HIGHLY RELEVANT TO YOUR STATED INFORMATIONAL NEED.

☐ WANT ONLY A FEW (I.E. A SAMPLE) OF THE DOCUMENT CITATIONS WHICH ARE HIGHLY RELEVANT TO YOUR STATED INFORMATIONAL NEED.

☐ WANT ONLY A FEW DOCUMENT CITATIONS REGARDLESS OF HOW RELEVANT TO YOUR STATED INFORMATIONAL NEED THEY MAY BE.

HITS: A ''HIT'' IS ANY DOCUMENT CITATION RETRIEVED BY THE COMPUTER WHICH IS RELEVANT TO YOUR INFORMATIONAL NEED.

HOW MANY ''HITS'' DO YOU FEEL YOU MIGHT NEED TO SATISFY YOUR INFORMATIONAL NEED? [CHECK ONE]

☐ 1 - 10 ☐ 11 - 20 ☐ 21 - 50 ☐ 50 OR MORE

~~~~~~

CLEARINGHOUSE USE ONLY:

☐ C ☐ E NO._____ CRBE: ☐ BIB ☐ 1 ☐ 2 ☐ 3 ☐ *
~~~~~~

NAME:_____. _____. DATE:____|____| '72

PLEASE INDICATE YOUR SATISFACTION WITH THE FOLLOWING BY CIRCLING ONE OF THE NUMERIC CHOICES GIVEN ON EACH SCALE.

A ''HIT'' IS ANY DOCUMENT CITATION RETRIEVED BY THE COMPUTER WHICH YOU DEEM RELEVANT TO YOUR INFORMATIONAL NEED.

HOW DO YOU FEEL ABOUT THE NUMBER OF HITS RETRIEVED BY YOUR SEARCH?

1	2	3	4
VERY DISSATISFIED	DISSATISFIED	SATISFIED	VERY SATISFIED

HOW DO YOU FEEL ABOUT THE OVERALL RELEVANCY OF THE HITS RETRIEVED BY YOUR SEARCH?

1	2	3	4
VERY DISSATISFIED	DISSATISFIED	SATISFIED	VERY SATISFIED

HOW DO YOU FEEL ABOUT THE AMOUNT OF TIME YOU SPENT (I.E. FROM THE TIME YOU ENTERED THE CLEARINGHOUSE UNTIL YOUR SEARCH STRATEGY WAS READY FOR SUBMISSION TO THE COMPUTER)?

1	2	3	4
VERY DISSATISFIED	DISSATISFIED	SATISFIED	VERY SATISFIED

HOW DO YOU FEEL ABOUT THE AMOUNT OF EFFORT THE CLEARINGHOUSE EXTENDED TOWARDS YOUR SEARCH NEED?

1	2	3	4
VERY DISSATISFIED	DISSATISFIED	SATISFIED	VERY SATISFIED

[PLEASE TURN TO NEXT PAGE.]

IN GENERAL, HOW DO YOU FEEL ABOUT THE RESULTS OF YOUR COMPUTER SEARCH EFFORT?

1	2	3	4
VERY DISSATISFIED	DISSATISFIED	SATISFIED	VERY SATISFIED

~~~~~~~~~~~~~~~~~~

HOW MANY HITS [NOT NECESSARILY TOTAL RETRIEVED CITATIONS] DID YOU RECEIVE? _____

HOW MANY CITATIONS DID YOU RECEIVE WHICH WERE NOT RELEVANT TO YOUR INFORMATIONAL NEED? _____

HOW MANY TOTAL CITATIONS [HITS AND "JUNK"] DID YOU RECEIVE?
_____

IF, AT ANOTHER TIME, YOU HAVE OTHER INFORMATIONAL NEEDS WHICH MAY BE SERVED BY ERIC, WILL YOU RETURN FOR ANOTHER COMPUTER SEARCH? ☐ YES ☐ NO

IF NO, WHY?_____

_____

WHAT IMPROVEMENTS WOULD YOU LIKE TO SEE MADE BY THE CLEARINGHOUSE TO MAKE COMPUTER SEARCHING OF ERIC FILES BETTER?

_____

_____

IF YOU WERE ADVISING A FRIEND CONCERNING THE COMPUTER SEARCHING SERVICES OFFERED AT THIS CLEARINGHOUSE:

WHAT WOULD BE THE STRONGEST POINT YOU WOULD DISCUSS?

_____

_____

WHAT WOULD BE THE WEAKEST POINT?

_____

~~~~~~

CLEARINGHOUSE USE ONLY:
☐ C ☐ E NO._____ TOT ____ ☐ KP ☐ VAR

Appendix D

Technical Notes

QUERY Software

The specific batch processing retrieval system used for the study was that provided by Computer Resources Corporation (CRC), 6825 Redmond Drive, McLean, Virginia, as modified by ERIC/CRESS. The CRC version of QUERY was F08APR70. The search parameters defined as minimum and used by both experimental groups were: (a) a maximum of five descriptors allowed, (b) no single descriptor searches allowed, (c) only the "or" and "and" search logic operators were allowed, (d) standard QUERY input formats were required, and (e) the TXT parameter was used throughout.

EDP Environment

Hardware. The following computer hardware, part of which was utilized during the study, was available at the NMSU Computer Center.

* IBM 360/65, 256K
* 2M bytes Ampex ECM (1.4 microsec interleaved)
* 2 Ampex disk spindles (30 microsec average access)
* 3 selector channels

  1 Mpx channel w/2 selector subchannels
* 1 1403 printer

The items marked with an asterisk (*) are critical capabilities but were not necessarily dedicated for the study. Additionally, it is known that the following reduced EDP environment will accommodate EQUIP and ERIC/QUERY system; however, substantially less response time may be expected.

IBM 360/50, 256K

1M bytes IBM LCS (8 microsec)

2 No. 2314 disk spindles (60 microsec average access)

2 selector channels

1 Mpx channel

1 1403 printer

System software. The NMSU computer utilizes IBM's OS version NVT, Release 19.6 in conjunction with HASP Version 3.0. The APL capability is IBM's Version 1, level 0 as modified by Dr. Tom H. Puckett. The APL workspace size is 35,000 bytes each. EQUIP consumes seven such workspaces for a total 245,000 bytes of which 210,000 are maintained on on-line storage devices (disk) at any given time. Dedicated active workspace within the Ampex ECM numbers to three before "swap out" occurs. The APL supervisor resides in IBM 360/65 core.

System work load. The NMSU multiprogramming system accommodates an average of 1,000 jobs per day--a situation which was not found to be detrimental to the study application.

704

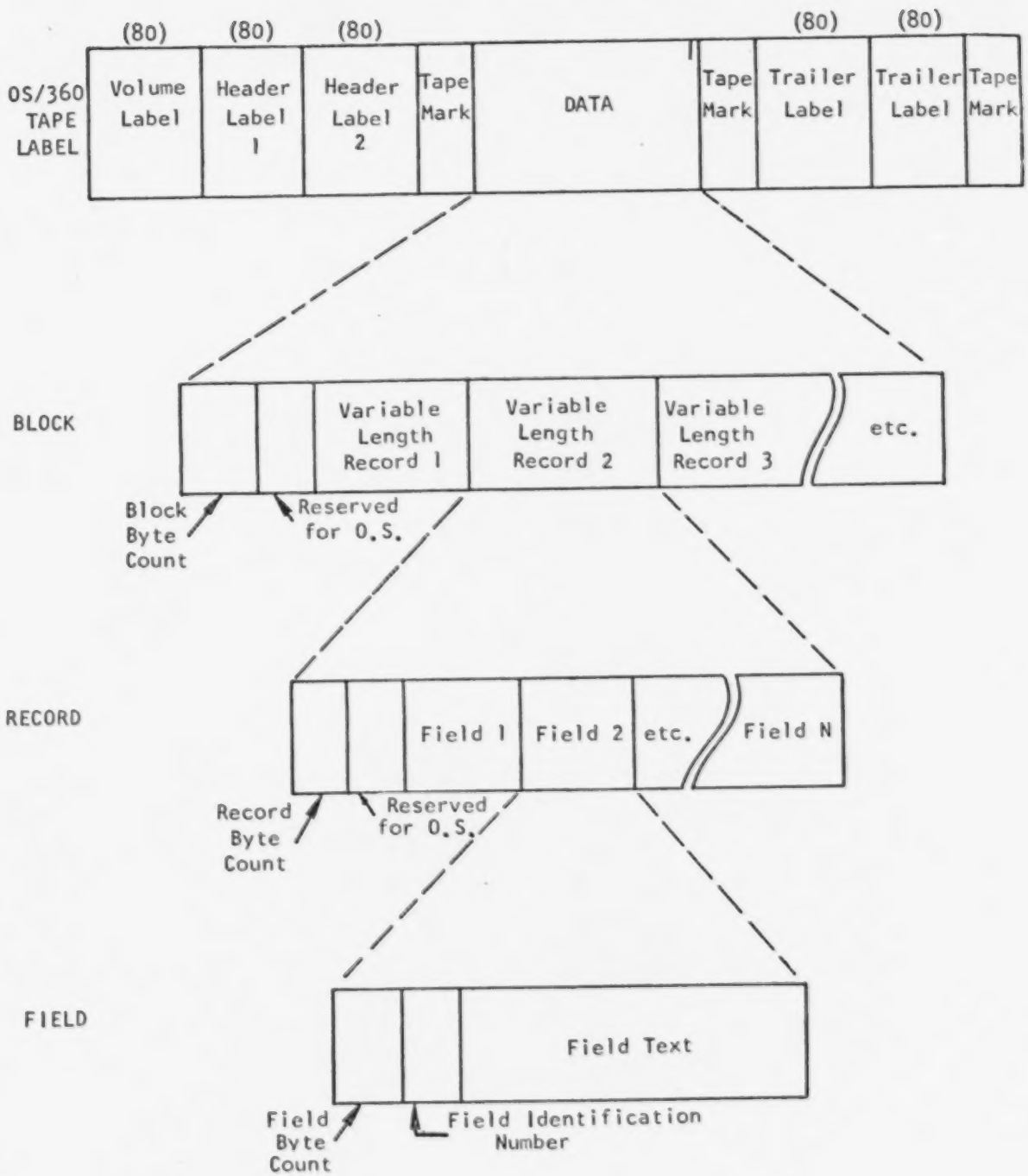
ERIC DATA ELEMENTS AND THEIR FIELD IDENTIFICATION NUMBERS

RESUME SUBSYSTEM

ORDERED BY FIELD IDENTIFICATION NUMBER

DATA ELEMENT/FIELD	KEYWORD	USED IN*	HEXADECIMAL	DECIMAL	NOTES
Sequence	None	F J P R(ALL)	00	0	
Add Date (Julian)	None	F J P R(ALL)	01	1	
Change Date (Julian)	None	F P R	02	2	
*Accession Number	ACC	F J P R(ALL)	10	16	
*Clearinghouse Acc. No.	CH	J R	11	17	
Other Accession Number	OCH	R	12	18	
Publication Type	PUBTYPE	R	13	19	OBSOLETE
Program Area	PA	P R	14	20	
Proposal Date	PROPDAT	P	15	21	
Project Date	PROJDAT	P	16	22	
Publication Date	PDAT	R	17	23	
Total Project Dollars	TOT	P	18	24	
Fiscal Year Funding	FYF	P	19	25	
Title	TITL	J P R	1A	26	
Personal Author	AUTH	J R	1B	27	
Institution	INST	F P R	1C	28	CODE #
Responsible Branch	BRANCH	P	1D	29	CODE #
Geographic Locality	GEO	F P	1E	30	CODE #
Cooperating Institution	CINST	P	1F	31	CODE #
Sponsoring Agency	SPON	P R	20	32	CODE #
Funding Agency	FUNDG	P	21	33	CODE #
Funded Individual	FUNDED	P	22	34	
Descriptors	DESC	F J P R(ALL)	23	35	
Identifiers	IDEN	F J P R(ALL)	24	36	
EDRS Price	PRICE	P R	25	37	
Descriptive Note	NOTE	R	26	38	
Project Officer	PRO	P	27	39	
Principal Investigator	PRI	P	28	40	
Field Reader	READ	F	29	41	
Address	ADDR	F	2A	42	
Issue	ISS	F J P R(ALL)	2B	43	
Abstract	ABST	F J P R(ALL)	2C	44	
Report Number	REPNO	R	2D	45	
Contract Number	CONT	F P R	2E	46	
Grant Number	GR	P R	2F	47	
Bureau Number	BN	P R	30	48	
Availability	AVAIL	J R	31	49	
Journal Citation	JNL	J R	32	50	
Experience Highlights	EH20-EH1	F	33-46	51-70	
Publications	PUBL	F	47	51	
Education	EDUC	F	48	72	
Region	REG	F	49	73	
Project Number	PN	P	5A	74	
Project Category	PCAT	P	5B	75	
Institution Name	None	F P R	80	128	28
Responsible Br. Name	None	P	81	129	29
Geographic Loc. Name	None	F P	82	130	(Derived 30
Cooperating Inst. Name	None	P	83	131	(From 31
Sponsoring Agency Name	None	P R	84	132	(Field 32
Funding Agency Name	None	P	85	133	33
Region Name	None	F	AD	173	73

F = Personnel, J = Journal Articles; P = Projects; R = Reports



* Numbers in parentheses indicate the length in bytes of fixed length data.